João Filipe dos Santos Sarraipa

Mestre em Engenharia Informática
Licenciado em Engenharia Electrotécnica e de Computadores

Semantic Adaptability for the Systems Interoperability

Dissertação para obtenção do Grau de Doutor em
Engenharia Electrotécnica e de Computadores

Orientador: Prof. Doutor Ricardo Luís Rosa Jardim Gonçalves, Professor Associado com Agregação da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa

Júri:

Presidente: Prof. Doutor Luís Manuel Camarinha de Matos
Arguente(s): Prof. Doutora Parisa Ghodous
Prof. Doutor Ricardo Jorge Silvério de Magalhães Machado

Vogais: Prof. Doutora Adina-Georgeta Bratu Cretan
Prof. Doutor Adolfo Sanchez Steiger Garção
Prof. Doutor António Carlos Bárbara Grilo
Prof. Doutor Hervé Panetto
Prof. Doutor João Pedro Mendonça de Assunção da Silva
Prof. Doutor Manuel Martins Barata
Prof. Doutor Rui Alexandre Nunes Neves da Silva

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Janeiro, 2013

À minha mulher Elsa.

Acknowledgments

I want to thank my supervisor: Professor Ricardo Jardim Gonçalves, that gave me opportunities, believing in my skills where others wouldn't took the risk and by encouraged me along these years to be better and better in the pursuit to be a good researcher.

Then, I would like to thank my colleagues at GRIS research group for their advices and ideas to the research work conducted that really contributed to this dissertation accomplishement: Pedro Maló; Carlos Agostinho; Fernando Ferreira; Hugo Vieira; Miguel Beça; José Ferreira; Carlos Coutinho; Edgar Silva; and Catarina Lucena.

I also want to thank to the MSc students that made their research in parallel to my research and which also contributed to the enrichment of my work: Diogo Figueiredo; David Gonçalves; Tiago Gaspar; Francisco Cavaco; Gonçalo Alves and Elsa Jesus.

A special thanks to João Mendonça by his friendly partnership in the research that together we have participated on.

Then I have to thank all my family and friends for their support. My parents, Jaime and Alice; my sisters Raquel and Lucilia as well their husbands José and Pedro and to my nephews: Carina; Stephanie; Diogo; Marisa; Michael and André, a special "thank you".

Also thanks to my jogging colleages Oliveira; Calado; and José Morgado that in our morning running sessions always pushed me to accomplish the thesis.

Last, but not least, I have to thank who I love and loves me and takes care of me every day: my wife, Elsa. She has always believed in me and her love has helped me to face the difficulties met in this last period.

Resumo

No contexto atual de negócio global e competitivo, é essencial que as empresas adaptem as suas bases de conhecimentos a fim de interagirem e colaborarem eficazmente. Contudo, devido à existente multiculturalidade de pessoas e empresas, existem diferentes formas de representação dos processos ou produtos de negócios, mesmo dentro de um mesmo domínio. Consequentemente, um dos principais problemas encontrados na interoperabilidade entre sistemas e aplicações de empresas está relacionado com a semântica. A integração e partilha do conhecimento empresarial para a construção de um léxico comum, desempenha um papel importante na adaptabilidade e integração semântica dos sistemas de informação. O autor propõe uma plataforma de trabalho que providencia uma estrutura de suporte ao desenvolvimento de sistemas para a gestão da dinâmica da adaptabilidade semântica. Esta permite que organizações distintas participem na construcão de uma base de conhecimento comum, deixando ao mesmo tempo que mantenham as suas próprias visões do domínio, sem comprometerem a integração entre elas. Assim, o sistema é capaz de estar atento a novo conhecimento, e ter a capacidade de aprender com ele, sendo capaz de administrar a sua interoperabilidade semântica duma forma dinâmica e adaptável. O autor também sublinha a visão que num futuro próximo, a capacidade da adaptabilidade semântica dos sistemas empresariais irá impulsionar a colaboração empresarial e o aparecimento de novos negócios.

Palavras Chave

Interoperabilidade Empresarial; Representação de Conhecimento; Interoperabilidade Semântica; Ontologias

Abstract

In the current global and competitive business context, it is essential that enterprises adapt their knowledge resources in order to smoothly interact and collaborate with others. However, due to the existent multiculturalism of people and enterprises, there are different representation views of business processes or products, even inside a same domain. Consequently, one of the main problems found in the interoperability between enterprise systems and applications is related to semantics. The integration and sharing of enterprises knowledge to build a common lexicon, plays an important role to the semantic adaptability of the information systems. The author proposes a framework to support the development of systems to manage dynamic semantic adaptability resolution. It allows different organisations to participate in a common knowledge base building, letting at the same time maintain their own views of the domain, without compromising the integration between them. Thus, systems are able to be aware of new knowledge, and have the capacity to learn from it and to manage its semantic interoperability in a dynamic and adaptable way. The author endorses the vision that in the near future, the semantic adaptability skills of the enterprise systems will be the booster to enterprises collaboration and the appearance of new business opportunities.

Keywords

Enterprise Interoperability; Knowledge Representation; Semantic Interoperability; Ontologies

Acronyms List

Acronym	Description
ABC	Activity Based Costing
ADDIE	Analysis, Design, Development, Implementation and Evaluation
AI	Artificial Inteligence
AIF	ATHENA Interoperability Framework
ANN	Artificial Neural Networks
AP236	Application Protocol For Furniture Catalog and Interior Design - ISO 10303-236:2006(E)
ATHENA	Advanced Technologies for Interoperability of Heterogeneous Enterprise Networks and their Application (research project acronym)
BIQMM	Business Model Measurement Interoperability Quotient
BN	Bayesian Network
CE	Concurrent Engineering
CKM	Collaborative Knowledge Management
CM	Communication Mediator
CPT	Conditional Probability Tables.
CRESCENDO	Collaborative and Robust Engineering using Simulation Capability Enabling Next Design Optimisation (research project acronym)
DAG	Directed Acyclic Graph
DIKUW	Data, Information, Knowledge, Understanding and Wisdom
DIKW	Data, Information, Knowledge and Wisdom
DSA	Dynamic Semantic Adaptability
EC	European Commission
EI	Enterprise Interoperability
EIF	European Interoperability Framework
EKT	Enhanced Knowledge Training
FCA	Formal Concept Analysis
FCW	Fundamentals in Collaborative Working
FEI	Framework for Enterprise Interoperability
FL	Fuzzy Logic

GC	Grand Challenge
GUI	Graphical User Interface
H1	Hypothesis 1
H2	Hypothesis 2
H3	Hypothesis 3
ICT	Information and Communication Technologies
ID	Instructional Design
IDEAS	Interoperability Development for Enterprise Applications and Software (research project acronym)
IDEF0	Integration Definition for Function Modelling
IEEE	Institute of Electrical and Electronics Engineers
IFAC	International Federation of Automatic Control
IIF	IDEAS Interoperability Framework
IMECE	International Mechanical Engineering Congress and Exposition
IMETI	International Multi-Conference on Engineering and Technological Innovation
INCOM	Information Control Problems in Manufacturing (conference acronym)
INNOVAFUN	INNOVAte FUrNiture business processes (research project acronym)
INTEROP	Interoperability Research for Networked Enterprises Applications and Software (research project acronym)
InterOP-VLab	The Virtual Laboratory for Enterprise Interoperability
IPB	International Product Business
IPR	Intellectual Property Rights
ISD	Instructional Systems Design
ISO	International Organisation for Standardisation (www.iso.ch - available at August 2011)
IST	Information Society Technologies (http://cordis.europa.eu/ist/ - available at August 2011)
ISU	Interoperability Service Utility
iSURF	An Interoperability Service Utility for Collaborative Supply Chain Planning across Multiple Domains Supported by RFID Devices (research project acronym)
IT	Information Technologies
ITPM	IT Portfolio Management
i-VLab	InterOP-VLab

KA	Knowledge Acquisition
KB	Knowledge Base
KMa	Knowledge Maintenance
KMo	Knowledge Modelling
KR	Knowledge Representation
KRE	Knowledge Representation Element
KRRM	Knowledge Representation Requirements Model
KT	Knowledge Training
KU	Knowledge Use
LCIM	Levels of Conceptual Interoperability Model
LMS	Learning Management System
LIMM	Language Independent Meta-Models
LO	Learning Object
LOM	Learning Object Metadata
MC	Markov Chains
MDA	Model Drivel Architectures
MLN	Markov Logic Networks
MENTOR	Methodology for Enterprise Reference Ontology Development
MMEI	Maturity Model of Enterprise Interoperability
MO	Mediator Ontology
MoMo	Model Morphisms
MOODLE	Modular Object Oriented Developmental Learning Environment
MSc	Master of Science
NG	Nominal Groups
NLP	Natural Language Processing
OL	Ontology Learning
oLEARCH	Ontology LEArns by seaRCHing
OWL	Web Ontology Language
PASCAL	Pattern Analysis, Statistical Modelling and Computational Learning
PhD	Doctor of Philosophy

PIER	Performance Indicators Evaluation Results
QICM	Qualitative Information Collection Methods
RQ1	Research Question 1
RQ2	Research Question 2
RQ3	Research Question 3
SECI	Socialisation, Externalisation, Combination, Internalisation
SIE	Semantic Interoperability Enhancer
SIRP	Semantic Interoperability Resolution Projects
SOKC	Spiral of Organisational Knowledge Creation
SMART-fm	A Standards compliant framework to support complete integrated product life-cycle information Management And electronic commerce foR the furniture manufacturing (FM) industry, in The advent of the smart enterprises – (research project acronym)
SME	Small and Medium Enterprises
STEP	Standard for the Exchange of Product model data
SyNAPSE	Systems of Neuromorphic Adaptive Plastic Scalable Electronics
TA	Training Analysis
TDes	Training Design
TDev	Training Development
TE	Training Execution
TI	Training Implementation
TiCW	Training in Collaborative Working
UPON	Unified Process for ONtology building
WWW	World Wide Web

Table of Contents

Table of Figures

Table of Tables

1 Introduction

In the current global and competitive context, for mostly of the enterprises, doing business globally has become critical to their survival. Additionally, enterprises, both big and small, need to establish cooperation agreements with others. Particularly, they have to combine forces to compete jointly in the market, in order to raise their own added value and to become specialised in niche activities. Thus, there is a demand for intelligent world-class solutions capable of reinforcing partnerships and collaborations.

However, enterprises are facing some difficulties concerning to the lack of interoperability of systems and software applications to manage and progress in their collaborative business [1][2]. But what Interoperability means? After a short study of the Interoperability word focused in its word formation it was reached three other (small) words: “inter, operate, and ability”. Where “inter-” is a prefix expressing “among, between” [3] or “cross over boundaries” [4]; “-ability” is a suffix expressing ability, fitness, or capacity [5], that in this case is related to the “operation” of something; and “operation” is related to “the act or process of operating or functioning” [6] that thing. Consequently, interoperability is the capacity of performing any kind of operation (could be through communication) across over certain boundaries. By other side, an interoperability problem is something that inhibits such operation ability.

Since the thesis is on the industrial information systems domain, the interoperability concept is presented on this document focused on the interoperation of systems. Thus, some of the reference definitions more used by the research community of this domain are:

- IEEE Dictionary defined interoperability as “the ability of two or more systems or components to exchange information and to use the information that has been exchanged” [7];
- Chen et al. defined that interoperability is considered as significant if the interactions can take place at least on three different levels: data, services and processes, with a semantics defined in a given business context [8][9];
- Panetto defined that interoperability is the ability of different types of computers, networks, operating systems, and applications to work together effectively, without prior communication, in order to exchange information in a useful and meaningful manner [10].

In 2003 Chen and Doumeingts stated that European Commission (EC) considers Interoperability of enterprise applications and software as a strategic issue. Moreover, compared with the situation in US or Japan, developing interoperability will bring more benefits to European industry in terms of knowledge sharing and collaboration, and thus accelerate European economy integration [11].

In accordance to such strategic issue the European Commission co-funded research IST projects that defined several frameworks related to this topic. A framework is a structure for supporting or enclosing something else, especially a skeletal support used as the basis for something being constructed [12]. An interoperability framework provides a set of assumptions, concepts, values and practices (methods

& tools) that constitutes a way of viewing and addressing interoperability issues [13]. In the following are presented some:

- Ideas Interoperability Framework (IIF)- The IDEAS[1] project (Interoperability Development for Enterprise Applications and Software, IST-2001-37368), identified the need for a structured approach to collect, identify and represent the current state of the art, vision statements, and research challenges in the area of Interoperability. It defined a framework for capturing and inter-relating this information from many perspectives called the IDEAS Interoperability Framework. This framework builds upon the meaningful achievement of interoperability between enterprises, which must be achieved on all enterprise layers (Figure 1.1) [9]:
    - Interoperability at business level should be seen as the organisational and operational ability of an enterprise to factually cooperate internally, but also, with other external organisations, whether these organisations are enterprises or public institutions.
    - Interoperability at knowledge level should be seen as the compatibility of the skills, competencies, and knowledge assets inside an enterprise and, with those of other, external organisations.
    - Interoperability at ICT (Information and Communications Technologies) systems level should be seen as the ability of an enterprise's ICT systems to cooperate internally and, with those of other, external organisations.
    - To overcome the semantic barrier, which emerges from different interpretations of syntactic descriptions, precise, computer processable meaning must be associated with each concept. It has to be ensured that semantics are exchangeable and based on a common understanding to be indeed a means to enhance interoperability.

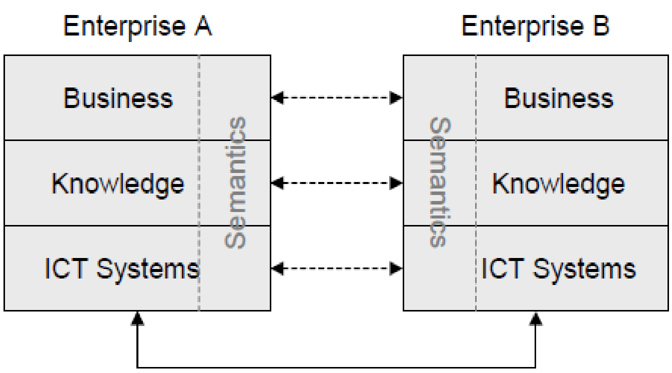


Figure 1.1 – Interoperability on all layers of the enterprise

- The European Interoperability Framework (EIF), in 2004, described interoperability in three aspects, which are, organisational, technical, and semantic. [14]:
    - **Organisational** – organising business processes and internal organisation structures

[1] IDEAS FP5 IST-2001-37368 - Thematic Network Interoperability Development for Enterprise Applications and Software - Roadmaps / Project duration: September 2002 to June 2003.

for better exchange of data.

- **Technical** – linking up computer systems by agreeing on standards for presenting, collecting, exchanging, processing, transporting data.

- **Semantic** – ensuring that transported data shares the same meaning for link-up systems.

- Interoperability Research Domains – The INTEROP[2] (Interoperability Research for Networked Enterprises Applications and Software) project aimed to extract value from the sustainable integration of the three key thematic components (that resulted from the roadmap for interoperability research made by the IDEAS project), to develop significant new knowledge [15]. Thus, INTEROP through a network of researchers develop research following three Interoperability Research Domains:

    - Ontology to identify interoperability semantics in the enterprise,

    - Enterprise Modelling to define interoperability requirements and to support solution implementation,

    - Architectures and Enabling technologies to provide implementation solutions.

- Athena Interoperability Framework (AIF) - The ATHENA[3] Interoperability Framework (AIF) defines an interoperability reference architecture that relates the modelling solutions coming from the three different research areas of ATHENA, namely enterprise modelling, architectures and platforms, and ontology. Figure 1.2 illustrates the reference architecture that focuses on the provided and required artefacts of two collaborating enterprises [13]. Interoperations can take place at the various levels:

    - **Interoperability at the enterprise/business level** should be seen as the organisational and operational ability of an enterprise to factually co-operate with other, external organisations in spite of e.g. different working practices, legislations, cultures and commercial approaches.

        - Related to this interoperability level appears the **Collaborative enterprise modelling** concept, **which** concerns to the exchange and alignment of knowledge models for describing the processes, organisations, products and systems in the collaboration context.

    - **Interoperability of processes** aims to make various processes work together. A process defines the sequence of the services (functions) according to some specific needs of a company.

[2] INTEROP-NoE FP6-IST 508011- Interoperability Research for Networked Enterprises Applications and Software: FP6 Network of Excellence from November 1st, 2003 to April 30th, 2007.
[3] ATHENA FP6- IST 507849 - Advanced Technologies for Interoperability of Heterogeneous Enterprise Networks and their Application: Integrated Project from 01/02/2004 to 31/03/2007.

- Under this topic was defined the Modelling of **cross-organisational business processes,** which focuses on defining process views that describes the interactions between two or more business entities.

- **Interoperability of services** is concerned with identifying, composing and executing various applications (designed and implemented independently). Services are an abstraction and an encapsulation of the functionality provided by an autonomous entity.

  - **Flexible execution and composition of services** is concerned with identifying, composing and executing various applications.

- **Interoperability of information/data** is related to the management, exchange and processing of different documents, messages and/or structures by different collaborating entities.

  - **Information interoperability** is related to management, exchange and processing of different documents, messages and other information structures.

- For each of the presented levels it was prescribed a **model-driven interoperability** approach where models are used to formalise and exchange the relevant provided and required artefacts that must be aligned and made compatible through negotiations and agreements. To overcome the semantic barriers which emerge from different interpretations of syntactic descriptions, precise, computer processable meaning must be associated with the models expressed on the different levels. It has to be ensured that **semantics** are exchangeable and based on common understanding in order to enhance interoperability. This can be achieved using **ontologies** and an annotation formalism for defining meaning in the exchanged models. The **model-driven interoperability** and the **semantics and ontologies** approaches to interoperability cut across the four levels and focus on integration of the corresponding interoperability approaches at these levels.

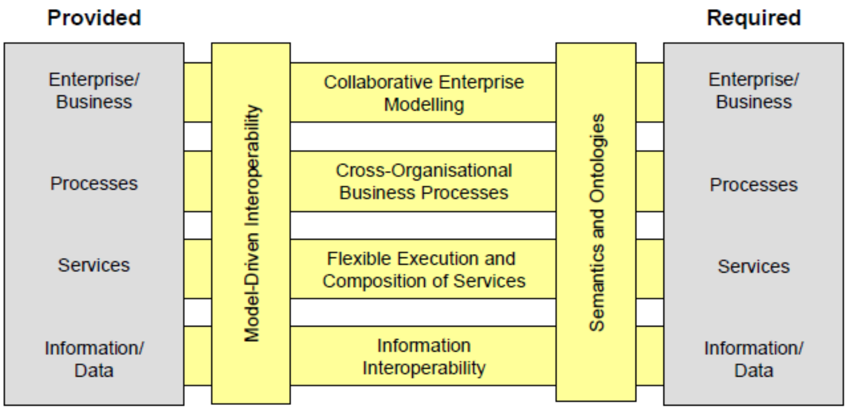


Figure 1.2 – ATHENA Interoperability reference architecture

The Interoperability research work that was supported and aligned to the presented frameworks naturally conducted to the appearance of the "**Enterprise Interoperability (EI)**" concept. This concept describes a field of activity with the aim to improve the manner in which enterprises, by means of ICT, interoperate with other enterprises, organisations, or with other business units of the same enterprise, in order to conduct their business [16]. Hence, the interoperability paradigm was co-related to the enterprises in the sense that they are the big stakeholders that could affect or be affected for such interoperability issues resolution.

Then it was defined a research roadmap related to EI, which purpose was to evaluate the EI developments on the past few years of the time that the roadmap was made (2008), and to interpret their impact in order to realign or fine-tune the EI long-term research goals [17]. Such study envisaged a vision that pursues what was stated in the i2010 communication[4], where its research key goal is to support new patterns of business that enhance innovation in enterprises and adaptation to new skill needs [18]. It also stated that European enterprises have to remain the principal benefactors of such research. This roadmap identified four Grand Challenges (GC) that represent the four global domains for research to reach the overall vision. These are:

- **GC1: Interoperability Service Utility**. This GC further extends the previous version, in the direction of focusing on the key services needed by the industry, including Small and Medium Enterprises (SME) specific needs and preconditions.

- **GC2: Future Internet and Enterprise Systems**. Extending the previous version GC2 (Web Technologies for EI), this GC sets the additional target of utilising and extending the new generation of enterprise systems, i.e. Enterprise Systems which are fully open, adaptive and integrated with innovation processes.

- **GC3: Knowledge-Oriented Collaboration and Semantic Interoperability**. This GC carries forward all the key topics from the previous version, further focusing on the needed semantics for organising, managing and exchanging knowledge and information – of both incoming and outgoing nature in the modern SME.

- **GC4: EI Science Base**. Keeping the main target of GC4, that is the empowerment of the scientific foundation of EI, new challenges are put forward as a result of the last 2 years of discussions among the EI community members in the area.

Although specific attempts for answering key challenges should be encouraged, the above GC are expected to be interrelated, when trying to provide complete, sustainable and adaptive solutions to the enterprise of tomorrow. As illustrated in the Figure 1.3, attempts towards the Interoperability Service Utility (ISU) are bound to drive and utilize developments in Future Internet and Enterprise Systems, while adopting and fertilizing evolutions in the Knowledge and Semantic Interoperability domain. A Scientific Base for EI should, in the meantime, act both as a repository of generated knowledge but

[4] Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions - "i2010 – A European Information Society for growth and employment" – June 2005.

also as an incubator of new ideas and future challenges to target.

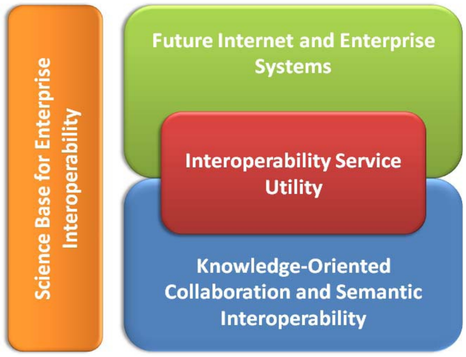


Figure 1.3 – GC of the Interoperability Research [17]

1.1 Motivation – The Semantic Interoperability Problem

The author participated in the mentioned projects (IDEAS; INTEROP and ATHENA) interacting directly as a researcher with a community that contributed to the evolution of the interoperability research on the past few years. Thus, due to this personal involvement on such research, and to author's background in engineering, with a tight relation to ICT, his motivation is naturally focused on the interoperability of industrial information systems. Regarding the topic industrial information systems, the author found a definition that addresses the specific topics he focused on his doctoral research:

> *The (*International Organisation for Standardisation) *ISO 16100-1 (2002) standard defines manufacturing software interoperability as the 'ability to share and exchange information using common syntax and semantics to meet an application-specific functional relationship through the use of a common interface' [19].*

Syntax is related to the rules, whereby words or other elements of sentence structure are combined to form grammatical sentences. Semantics comes from the ancient Greek for 'meaning'. The common way to specify meaning is by giving a definition. The international standard ISO 1087 tells that a 'definition' is a "Verbal description of a concept, permitting its differentiation from other concepts within a system of concepts" [20]. A definition can be given either in informal or formal language. Informal semantics is about creating meanings that people can understand; and formal semantics is about creating meanings that machines can process. From these descriptions, semantics is an interesting topic with potential to be considered as the main focus of the research to be followed.

However, in additional, the author analysed the semantics role and importance on the previous presented European research initiatives. All of them presented semantics has one of the important fields or issues to be addressed on the interoperability research domain.

- IIF presented semantics as a vertical issue crossing all the other topics: business; knowledge and ICT systems.
- EIF considered semantic as one of its three mains aspects defining interoperability.
- INTEROP defined one of its three domains as responsible to identify interoperability

semantics in enterprises. This domain is ontology.

- AIF as a merging result of IIF with INTEROP interoperability specifications defined semantics and ontologies as a vertical issue/characteristic that crosses all the interoperability desired levels: enterprise/business; processes; services and information/data.

- Finally, EI roadmap established for one of its GC the knowledge oriented collaboration and semantics interoperability. This GC focuses on the needed semantics for organising, managing and exchanging knowledge and information in SMEs.

Consequently, the importance that researchers community has given to semantics on the interoperability research domain, helped author to decide to be involved on such interoperability domain.

1.1.1 Research Objective to Follow

Semantic interoperability concerns the ability of two systems to present and process, user information tasks in a way that meanings of those tasks are retained [21]. To depict it, in the *Figure 1.4*, the left enterprise only understands triangles and the right enterprise only understands circles. To achieve semantic interoperability both enterprises should understand each other. On *Figure 1.4* a) is foreseen a solution, which "translates" circles in triangles and vice-versa solving this interoperability semantic problem.

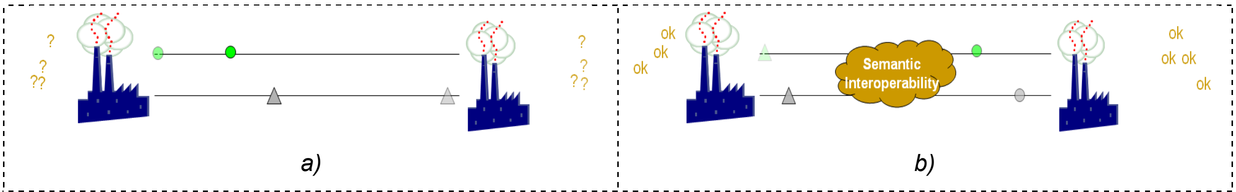


Figure 1.4 – Semantic Interoperability in communications

In alignment with this statement, the author made a first background analysis related to this issue to evaluate what kind of requirements could appear on defining solutions to the semantic interoperability issues. Thus, in his involvement in the INNOVAFUN project[5], a project developed in cooperation with the funStep initiative group[6], which intends to facilitate interoperability in the exchange of furniture catalogues and information between furniture stakeholders, the author contributed with a definition of a use case. It is named International Product Business (IPB), since its aim is to represent situations, which could facilitate the description of how the knowledge is handled in an international plan of the furniture business. This use case is presented further on this document (sub-chapter 5.1.1). Such work conducted the author to research situations related to semantic interoperability and multilinguism or regional standards, which reached a common understanding among developers, system users, and domain experts in the furniture domain. On that work, as conclusion, it was identified the need of having a structured furniture Knowledge Representation (KR) able to facilitate the full information

[5] European project INNOVAFUN "Applying open standards to INNOVAte FUrNiture business processes" - http://standards.eu-innova.org/Pages/InnovaFun/default.aspx.
[6] funStep Interest Group - a group focused in helping furniture business stakeholders in solving software issues related to its business information exchange and promoting its solutions. http://www.funstep.org

interoperability establishment in the furniture business. Thus, semantics has a distinguished role in the interoperability area and as by heritance on EI.

In addition, it was verified that from a business point of view, enterprise's competitiveness is to a large extent determined by its ability to seamlessly interoperate in a collaboratively way with others. Consequently, to meet such collaborative business objectives, enterprises need to adapt their knowledge resources in order to smoothly interact with others. The smooth adaptation of the enterprises knowledge resources requests an advanced characteristic to the interoperation that is dynamicity. This originates the creation of the **semantic adaptability** concept. This concept represents the capacity level of adapting the existent semantics to new acquired knowledge by a system or actor. The dynamic characteristics of the systems require a high level of semantic adaptability and as consequence intelligence, which is an important characteristic to be addressed on this research work. All these statements justify the author's thesis research focus that resulted in his thesis title: **Semantic adaptability for the Systems Interoperability**. It means that author's main objective is to contribute to the systems interoperability with a novel proposal for supporting solutions and characterizations related to the semantic adaptability, which is aligned with the following statement:

Nowadays, due to the proliferation of terminology, organizations from similar business environments have trouble cooperating, and are experiencing difficulties exchanging electronically vital information, such as product data, even when using international standards. Consequently there is a demand for intelligent world-class solutions capable of reinforcing partnerships and collaborations with an improved cross-cultural understanding [22].

1.2 Research Method and Approach

Research is a systematic process of collecting and analysing information to increase our understanding of the phenomenon under study. It is the function of the researcher to contribute to the understanding of the phenomenon and to communicate that understanding to others [23].

Figure 1.5 presents the research method adopted by the author, which is an instantiation of the 7 classical phases of the scientific method [24]. The difference is that this has 8 phases and has a cycle that encloses all the phases. It means that for reaching the overall objective, which in this case is related to an entire Doctor of Philosophy (PhD) thesis, all the phases will be conducted several times. Apart of having a set of main research questions related to the thesis research goal, there will be several sets of "small" research questions for each cycle that will represent a small research step forward on the track to reach the overall result. Each of these cycles could end by the accomplishment of a publication about a specific topic, and sometimes in a technology transfer to industry. Consequently until the publishing of the thesis at phase 7, the author has to conduct the work accordingly to all the classical phases of the scientific method the number of times necessary to mature his thesis global idea.

In this process, scientific results will appear increasingly on time, accordingly to the knowledge maturation of the researcher on the topic. Such scientific results intend to contribute to the science frontiers advance and, if not effectively, at least to potentiate, technology transfer to industry.

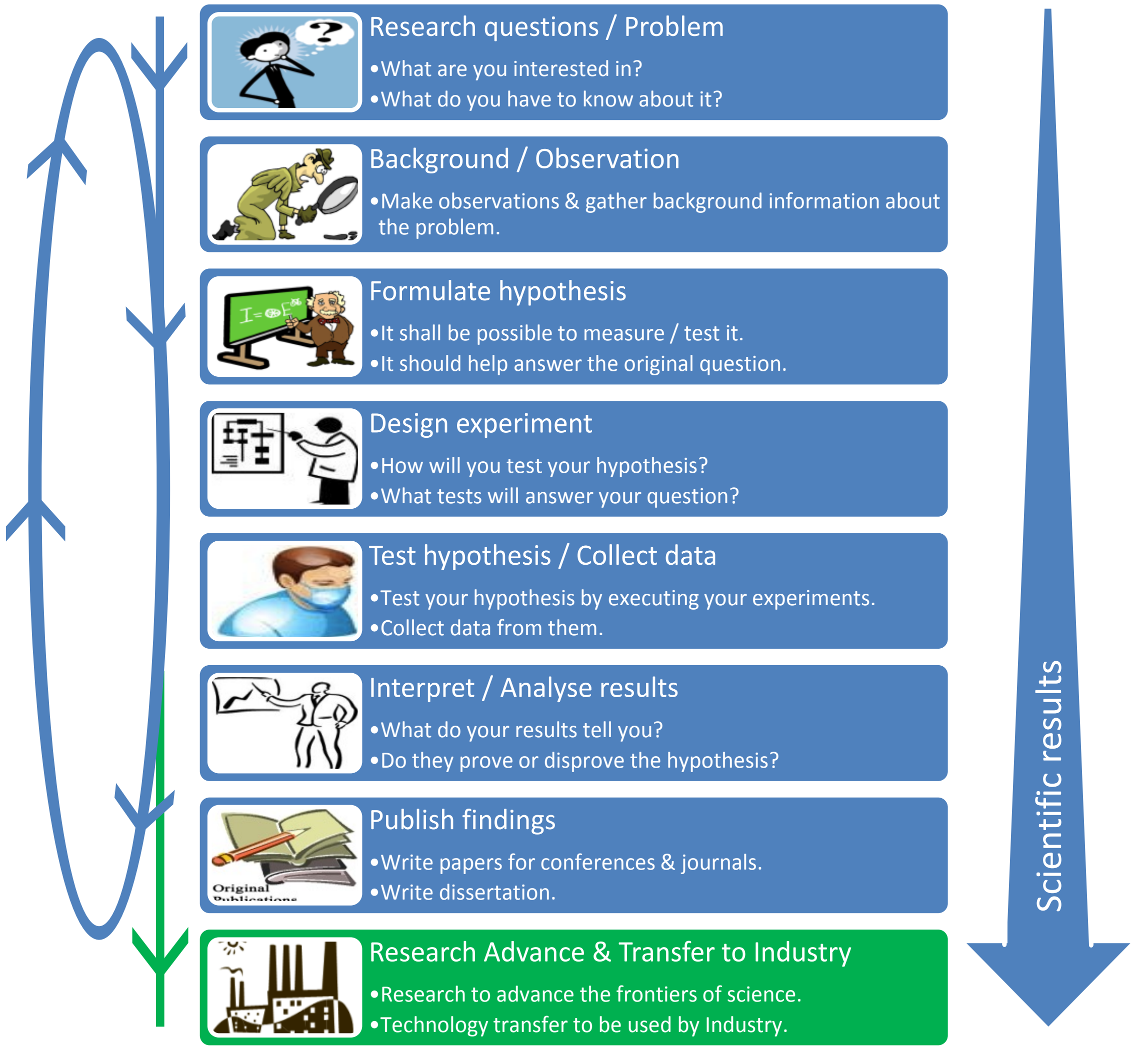


Figure 1.5 – Adopted Research Method

This document presents all the classical phases of the research method. It presents in the first chapter, the research questions to which the author is interested in, supported by an initial background observation that afterwards was used to formulate the hypothesis to answer the presented research questions. Then, in the second chapter, it presents a literature review to better understand the research area and, to accomplish the background observation needed, to properly design the experiment at the chapter 3, which in this dissertation is related to a framework for dynamic semantic adaptability resolution. The following four chapters (4; 5; 6 and 7) are related to the hypothesis tests and collection of data, whose present proof of concepts prototypes of the framework presented previously. Finally in the last two chapters are analysed all the research results in order to know if the approach conducted proves the thesis and solve the research questions identified. This entire document, by itself, represents a publishing finding, which is complemented by all the publication in conferences and journals, which aggregates the scientific results that the author made in his overall PhD work. In addition, the chapter number 8 and 9 of this document present an analysis of the research conducted, which describes the contribution for the research advance and technology transfer to industry.

1.2.1 Research Questions

The research questions addresses the focus on the author intends to participate and contribute for. Thus, the questions to be addressed by this thesis are:

- **RQ1 (Research Question 1):** How to improve the interoperability of software applications and information systems semantics?

- **RQ2 (Research Question 2):** What is the role and implications of the semantic mapping in a network of information systems and software applications?

- **RQ3 (Research Question 3):** Can the harmonisation of KR contribute to the advance of the integrated information system's intelligence?

To contribute to respond to this research questions it is needed to make a first background observation to identify some hypothesis in their resolution.

1.2.2 First Background Observation

Currently, it is widely accepted that systems that possess knowledge and are capable of decision making and reasoning are regarded as 'intelligent' [25][26]. There are recognised techniques, such as Fuzzy Logic (FL), Artificial Neural Networks (ANN), machine learning and evolutionary algorithms that contribute to increase a system's 'machine intelligence quotient' [27]. The rationale behind the intelligent label of those techniques is their ability to represent and deal with knowledge [28]. These branches form the triad of the so-called computational intelligence [29].

Ontologies play an important role in intelligent systems. An ontology is an explicit specification of a conceptualisation [30] that refers to the shared understanding of some domain interest, which may be used as a unifying framework to facilitate knowledge sharing and interoperability between independently developed subsystems [31]. Thus, ontologies allow key concepts and terms relevant to a given domain to be identified and defined in a structure able to express the knowledge of an organisation. Its recognised capacity to represent knowledge, to facilitate reasoning, use and exchange knowledge between systems contributes to increase the computational intelligence.

Semantic resolutions are related to the ability to manage concepts meanings and consequently its knowledge. Thus having information systems with some intelligence able to acquire knowledge and maintain it dynamically to establish interoperable communications with others is the objective to follow.

1.2.3 Aimed Contribution

The aimed contribution is to contribute with methodologies & approaches to solve semantic interoperability problems in the enterprises information systems. Since semantics is directly associated to the knowledge, the author structured his research work accordingly to the knowledge life-cycle stages. Thus, the main idea has been to contribute with possible technical solutions to each of such stages, providing also a conceptual contribution on the addressed concepts.

The objective is to identify a new methodology able to facilitate knowledge acquisition and organisation that could represent a set of enterprises in order to be the mediator in their communications without forgetting the dynamics of such system. Thus, he will identify a way to provide such methodology with Ontology Learning (OL) capabilities focused on users' interactions. The real outcome expected is to contribute with solutions to the system's intelligence increase in the area of semantic adaptability of systems. All of these research actions will follow the overall EI roadmap directives.

1.2.4 **Hypothesis**

The hypotheses considered to execute experiments are the following:

- **H1 (Hypothesis 1):** A proper interoperability system based on KR and reasoning is able to be adapted based on external feedback, facilitating the semantic adaptability on future enterprise systems;

- **H2 (Hypothesis 2):** With a proper methodology for harmonisation of ontologies, the semantic interoperability in an network of enterprises is facilitated;

- **H3 (Hypothesis 3):** The next generation of intelligent systems to assist on interoperability of software applications and information systems needs the support of machine learning and operational research methods.

1.3 Dissertation Outline

Nine chapters, plus an annexes and a references section compose this dissertation. The first chapter introduces the dissertation by summarizing the extensive research work done, through various European initiatives, since a decade, in the domain of systems interoperability. Some definitions and frameworks are presented to explain the conception of the EI concept. In additional to it, a roadmap is presented with the goal of presenting the actual stage of the research on the domain. Then, author presents a short explanation to describe his motivation to the semantics interoperability research area, accomplished by his research objective definition. Afterwards, it is presented the research method followed on this dissertation, which is followed by the research questions definition and a first background observation description. As a result of such first background research it is presented the author's aimed contribution to the problem defined, and finally the hypothesis to be followed on the dissertation resolution are described. The chapter finishes by presenting the dissertation outline.

The second chapter is an extensive literature review on KR and Knowledge Modelling (KMo) techniques and tools, which is organised in a structure inspired in the Nonaka and Takeuchi knowledge creation model. Thus, its main sections are: Knowledge Acquisition (KA); KMo; Knowledge Use (KU); Knowledge Maintenance (KMa) and Knowledge Training (KT). KA section presents techniques and solutions for knowledge extraction or elicitation from tacit or explicit sources. KMo section presents various concepts, methods and tools related to the purpose of transforming knowledge into reusable and explicit format and then establishing combinations for its preserving, improving, sharing, aggregating and processing. KU section introduces a formalisation of a tuple for the use of the stored knowledge to facilitate interoperability and to potentiate semantic adaptability of the systems. KMa section introduces theories techniques and methods related to the Artificial

Intelligence (AI) domain able to provide learning mechanisms for automatic knowledge updates accordingly to information systems evolution. KT introduces training concepts, tools and methodologies for training development to implement a knowledge transfer solution with the purpose to deliver knowledge to help competences and skills development on key people to handle systems or tools for semantic interoperability resolution. All of these topic sections are accomplished by conclusions that intend to provide guidelines to the research and prototype solutions developments.

The third chapter presents the author's conceptual contribution to this dissertation. It presents a framework that intends to represent a structural basis for supporting the development of systems or tools for dynamic semantic adaptation resolution. Such framework is based on two abstract systems, whose follow the view that a system able to handle knowledge understandability, it is able to manage its semantics. Thus, these two abstract systems were designed in a direct relation to the various knowledge life-cycle phases. In addition, epistemological dimensions of knowledge are introduced to give some assumptions about the dynamic capacity level that systems are nowadays able to reach.

The fourth chapter presents a methodology for enterprise reference ontology development. Its main objective is to facilitate a collaborative building of an ontology to be the reference for a group of enterprises enabling at the same time these enterprises keep their original information model unchanged. This is possible due to the available functionality for mapping tables' establishment between those information models and the reference ontology. Thus, it describes a set of steps to be followed in a shared ontology building for traceability of semantic evolutions or changes and semantic operations used to basic lexicon and ontologies mappings establishment. It also presents a prototype tool related to this methodology.

The chapter number five presents a knowledge organisation architecture for multicultural scenarios, which was defined and tested for the manufacturing systems of the furniture industry. It encloses in a single architecture several KREs as dictionaries, thesaurus, ontologies and data standards to handle the semantic enrichment of a standard to be used in the furniture catalogues categorisation and in its supply chain communications.

The chapter number six presents a tool, which is based on the premise that ontologies "learn by searching". It is a searching products tool that combines semantic mappings with user's feedback, providing an automatic learning to ontologies enabling auto-adaptability to the semantics of the information systems. The feedback given by the users are used to get some patterns. These patterns are analysed through machine learning algorithms accomplished with statistics, which enables ontologies to learn by readjusting or updating their concepts weight readjustment or by introducing new concepts. This achieves the dynamic capacity of handling semantic of information systems. A prototype related to this is also demonstrated.

Chapter 7 presents a methodology for a simple and organised way of developing training materials for courses. The implementation of a customisable e-Training service that uses an ontology to handle the training knowledge is also introduced. Both approaches enable the automatic orchestration of courses according to specific user needs or profiles, facilitating the knowledge transfer to people. In additional, it implements flexibility in e-Training development, deployment and reuse of training materials.

In chapter 8 it is analysed the research work conducted, checking if the research questions identified by the author, were answered following the proposed hypothesis and finally if the thesis was answered. However, in this chapter it is also presented a semantic checking framework and a business validation methodology, which objective is to contribute with two solutions for semantic interoperability resolution projects evaluation that were also used in this research. In additional, this chapter presents an analysis related to this dissertation research positioning in comparation to other research works in the same domain. Such analysis in conjunction to others were then used to accomplish an overall validation of the research conducted in relation to: research organisation and integration; scientific position y publications; industrial and technical developments or results.

Chapter 9 presents some final considerations about the research results achieved enphasizing the real added value of it to the research community. A future work description is also introduced.

The document ends with a section that presents some annexes followed by the bibliographic list of references read for the development of this work.

2 Literature Review

With the current economic downturn, margins are being squeezed in most industries, resulting in an urgent need for higher efficiency and effectiveness, including the putting in place of better-value ways of working. A true e-business of the twenty-first century is one which thinks and acts in a way that allows it to collaborate, integrate and empower by: a) internal and external business processes working together seamlessly, enabling collaboration with suppliers, partners, employees and customers across usual enterprise boundaries; and b) ensuring that employees have at their fingertips the knowledge, applications and services needed for the job [32].

Thus, organizations need to be interoperable to achieve such seamless cross-organizational business interaction and must also develop collaborative working practices as to improve productivity and to take an holistic approach to implementing product phases with emphasis e.g. on applications of problem solving, creativity, participatory and knowledge-based design and engineering [33]. This approach to engineering increases the flexibility and responsiveness of the organization is presence of complex and varying customer demands [34].

To answer these challenges of globally distributed organizations facing constant transformation of work, means of ICT supported inter-organisational knowledge-creation is needed to facilitate the emergence of business global networks. In these networks, knowledge is a critical resource for development, and a fundamental challenge is to organize work with knowledge in a way that facilitates continuous knowledge advancement and supports the sharing of intellectual achievements among the members of the community [35].

Consequently, it is necessary to better understand the processes of knowledge creation. Philosophy has investigated knowledge creation for millennia, but has concentrated in the last fifty years on macro-theories of knowledge creation on a grand historical scale; the knowledge economy, on the other hand, requires micro-theories of knowledge creation applicable to today ICT system's needs. Therefore, many new micro-theories of knowledge and technology creation have emerged in the last decade of the 20th Century, from fields outside of philosophy [36], to answer to the crucial necessity of sharing, transferring, and creating knowledge in the today dynamic markets.

The choice of a knowledge creation model is crucial to the guidance of this dissertation research, since it will determine what fields should be addressed. The author has used the Harzing's tool called Publish or Perish[7] to help on his research citation analysis. It provides a good search facility enabling the researcher to found the most important publications in a specific matter. On the case to knowledge creation model, it was clearly identified the one defined by Nonaka and Takeuchi [37], as the most referred work and with a higher impact related to the micro-theories of knowledge creation in the ICT area. Moreover, due to its importance on the field, most of the researchers that introduced new micro-theories analysed their work in relation or as an extension to the Nonaka and Takeuchi's model. Paavola et al. [38] made a comparison between three influential models concerning learning and

[7] Harzing - Research in International Management Products & Services for Academics http://www.harzing.com/pop.htm#whatfor

innovative inquiry, which appear to represent essential aspects of the knowledge creation process, whose are, as they consider, the model of knowledge building by Bereiter [39], the model of knowledge creation by Nonaka and Takeuchi [37], and the model of expansive learning by Engeström [40][41]. The author presents here this analysis in the sense that it helps to introduce these models main characteristics. Paavola et al. [38] established a comparison of such models, following a schematic representation of the relations between them (Table 2.1).

	Nonaka & Takeuchi	***Engeström***	***Bereiter***
The role of individual expertise	*Black box, individuals create knowledge*	*Socially embedded*	*Theory of expertise*
Main focus	*Tacit knowledge (insighting)*	*Knowledge embedded in practices (acting)*	*Knowledge objects (conceptualizing)*
Type of processes focused	*Emphasize bodily processes, personal experience*	*Emphasize material object-oriented activities*	*Emphasize solving of knowledge problems*
Source of innovation	*Transforming tacit knowledge to explicit knowledge*	*Overcoming tensions, disturbances, and ambiguities by expansive learning*	*Working deliberately for extending and creating new knowledge objects*
Scope of framework	*Different ontological levels (individual, innovative team, organization, and inter-organization level)*	*Activity systems and networks of activity systems*	*Knowledge-building communities*
Educational application	*Knowledge-creating schools*	*Expansive learning school*	*Schools as knowledge building communities*

Table 2.1 – Schematic Representation of Nonaka & Takeuchi; Engeström and Bereiter knowledge models comparison [38]

From the presented table it could be concluded that Engeström focuses his model on the knowledge related to the practices (acting). Bereiter by other side is more related to processes for solving knowledge problems enclosing communities. Nonaka and Takeuchi’s model focuses more on individuals’ personal knowledge and in transforming it to explicit knowledge.

Most companies invest in their knowledge assets by recruiting knowledgeable people in the first instance and then further by training them. The challenge for organisations is how to keep the knowledge and information they have invested in. 'Knowledge' is not all held in data capture systems; much of it is held within people. This means that knowledge and people are inextricably linked [42], thus knowledge management needs to be perceived and understood as the dynamic assembly of information (explicit) out of a specific context, merged with an individual’s personal knowledge (tacit) captured, internalized and applied e.g. in decision processes/methods [43]. These designated (value added) methods are enterprise's "knowledge assets", whose are the delivery vehicles for increasing the velocity and flexibility necessary to please customers, beat competitors and ensure success in increasing enterprise competitiveness[44].

Since the dissertation is related to semantics issues resolution on enterprise systems to facilitate business collaborations, it implicates an appropriate enterprise knowledge assets management that requires a dynamic and flexible enterprises knowledge explicit and tacit handling. Thus, Nonaka's and Takeuchi's model prevails as the choice to follow on the research conducted. However, some other models characteristics should be addressed e.g. the conceptualization of knowledge objects (from Bereiter model) or the frameworks whose are related to systems and networks of activity systems as addressed by Engeström model.

Consequently, the following sub-sections proceeds with the presentation of the literature review categorised by the various knowledge life-cycle plus the "knowledge training" stages, accordingly to its direct relation with Nonaka and Takeuchi's model, as detailed in chapter 3.

2.1 Knowledge Acquisition (KA)

Knowledge is defined by the Oxford English Dictionary as (i) expertise, and skills acquired by a person through experience or education; the theoretical or practical understanding of a subject, (ii) what is known in a particular field or in total; facts and information or (iii) awareness or familiarity gained by experience of a fact or situation. The term *knowledge* is also used to mean the confident understanding of a subject with the ability to use it for a specific purpose if appropriate [45]. By other side, philosophical debates in general start with Plato's[8] formulation of knowledge as a "justified true belief" that states that *in order to know that a given proposition is true, one must not only believe the relevant true proposition, but one must also have justification for doing so [47]*.

When someone "memorizes" information (as less-aspiring test-bound students often do), then they have amassed knowledge. This knowledge has useful meaning to them, but it does not provide for, in and of itself, an integration such as would infer further knowledge [48]. This is not the real purpose of knowledge. Thus, knowledge is a deterministic process. In additional, knowledge pursues the gathering of new knowledge in a kind of not ending cycle. KA is the action beyond such process. KA involves complex cognitive processes: perception, learning, communication, association and reasoning. However, in Knowledge engineering, KA includes the elicitation, collection, analysis, modelling and validation of knowledge [49]. Even so, the full automatic machine KA remains in the distant future [50][51], resulting in a process extremely expensive [52][53]. Since this dissertation is following in separated each of the knowledge life-cycle stages, KA is reduced for the elicitation of knowledge, putting other characteristics discussions for other appropriated stages' sub-chapters. The advanced KA, as the philosophical understanding is far from being feasible by an information system, thus, KA main objective is reduced here, to the elicitation of knowledge from explicit and tacit sources.

Knowledge elicitation is one of the most difficult and error-prone tasks that a knowledge engineer does

[8] Plato (428/427 BC – 348/347 BC), was a Classical Greek philosopher, mathematician, writer of philosophical dialogues, and founder of the Academy in Athens, the first institution of higher learning in the Western world. His mentor was Socrates, and his student was Aristotle [46].

while building a knowledge-based system. The cost and performance of a Knowledge Base (KB) building depends directly on the quality of the knowledge acquired. During this process it is needed to identify where in the organization the knowledge exists and how to capture it [54].

2.1.1 Explicit Knowledge Elicitation

The KA is a bottleneck of the Artificial Inteligence (AI). How to extract or elicit knowledge effectively from the text is what actually, knowledge engineering research is mostly concerned. Such ability with the identification of various types of semantic relationships between the terms from unstructured text is considered to be of critical importance on this research domain [55][56][57]. There is several research initiatives on the area like PASCAL[9], which is a Network of Excellence funded by the European Commission. PASCAL is developing the expertise and scientific results that will help create new technologies such as intelligent interfaces and adaptive cognitive systems. To achieve such, it supports and encourages collaboration between experts in machine learning, statistics and optimization, to promote the use of machine learning in many relevant application domains such as: machine vision; speech; haptics; brain-computer interface; user-modelling for computer human interaction; multimodal integration; Natural Language Processing (NLP); information retrieval; and textual information access. Some of these mentioned domains, which PASCAL addresses are not considered on this dissertation (e.g. speech; haptics; etc.), would be related to the tacit knowledge elicitation and not explicit.

As a result to these kind of initiatives, many knowledge acquisition techniques and methods have been studied, and various NLP tools have been introduced. One of these solutions is the *TermExtractor* - a web application able to learn and extract terminology related to a specific domain from text [58].

Other existent kind of solutions is related to OL from text, which intends to be an automated construction and population of ontologies. A specific work example related to that, has been conducted by Cimiano et al. in [59]. They defined an approach for the automatic acquisition of taxonomies or concept hierarchies from a text corpus based on Formal Concept Analysis (FCA). This is a method mainly used for the analysis of data, i.e. for investigating and processing explicitly given information [59].

These entire NLP approaches use machine learning techniques, as Latent Semantic Analysis (LSA)[10] or Latent Semantic Indexing (LSI) algorithms to extract and classify terms based on their role on phrases and taking in consideration some statistics patterns of their distribution between corpora.

However, concepts structuring that involve understanding of semantics and contexts still be handled by a knowledge engineer. Each approach could be considered automatic, but as Cimiano et al. stated: “we do not believe in fully automatic ontology construction without any user involvement. In this sense, in the future we will explore how users can be involved in the process by presenting him/her

[9] Pattern Analysis, Statistical Modelling and Computational Learning (PASCAL) – Accessed at August 2011: http://www.pascal-network.org/

[10] Latent Semantic Analysis (LSA) is a theory and method for extracting and representing the contextual-usage meaning of words by statistical computations applied to a large corpus of text [60]

ontological relations for validation in such way that the necessary user feedback is kept at a minimum [59]".

2.1.2 **Tacit Knowledge Elicitation**

Researchers have highlighted the importance of eliciting knowledge from individuals and organisation as a mechanism for preservation knowledge and experience, improvement of knowledge reuse, and acceleration of processes such as individual and organisational learning. In additional, such tacit elicitation is also important for other situations, as for special meetings where knowledge come from discussions of different points of view of several persons. Many techniques have been developed to help elicit tacit knowledge from an expert, like the protocol-generation type, which aggregates interviews, reporting, and observation. But there are many others like concept mapping, interviews, knowledge audits, cognitive modelling, data analysis and work patterns analysis.

If it is taken in consideration the elicitation that aggregates several domain engineers' participants, whose are collaboratively discussing different views, as for the agreement of meanings of concepts, processes of work and others, what is important is how to organise the discussions. It was chosen this kind of elicitation case, mainly because this dissertation is focused in addressing the problems of semantics between enterprises, thus there is mostly a need of establishing a common view of such domain semantics. In additional, according to Hoffman *et al*. (1999) one of the most important approaches to knowledge elicitation is to use a combination of existing methods according to the conditions of the organisation being analysed [61][62].

Thus, to capture what people say on such meetings about their meanings and interpretations, it is needed to use Qualitative Information Collection Methods (QICM). Qualitative since it typically involves qualitative data, i.e., data obtained through methods such interviews, on-site observations, and focus groups that is in narrative rather than numerical form [63]. By other side quantitative inquiries use numerical and statistical processes to answer specific questions, where statistics are used in a variety of ways to support inquiry or program assessment/evaluation [63], which is not the purpose of this case.

As a conclusion, the employ of QICM process would effectively improve the approach to elicit knowledge from domain experts in the mentioned situations. Most known QICM identified by the author are presented in the following sub sections.

Nominal Groups

A Nominal Groups (NG) session requires several participants to discuss about a topic to generate a list of ideas. This approach is an alternative to brain storming, because it is a more structured discussing method. In this method are given time to participants to think and write down their ideas before telling them to the group [64]. This technique is a structured variation of small group discussion methods. The process prevents the domination of discussion by a single person, encourages the more passive group members to participate, and results in a set of prioritized solutions or recommendations. The steps to follow in this technique are [65]: 1) silent generation of ideas in writing; 2) round robin recording of ideas; 3) serial discussion for clarification (for each idea); 4) voting on the priority strategies; 5) discussion of preliminary voting; and 6) final voting.

In addition, NG method has many benefits and advantages that include: 1) motivates all participants to get involved; 2) generates many ideas in a short period of time; 3) obtain ideas from people of different backgrounds and experiences; 4) stimulates creative thinking and effective dialogue; 5) allows clarification of ideas.

The relevant disadvantages can be related to: 1) the necessity of having a skilled leader; 2) it can take too much time if the group is not properly controlled and is allowed to run for too long; and 3) assertive personalities may dominate unless leadership skills are exercised.

Metaplan

The Metaplan technique is a learning method especially for groups. It is a collaborative and moderated technique with focus on solving group decisions or problems [66]. It combines individual and collective contributions and is used to organize concrete ideas into more general conclusion leading to recommendations. It is mostly useful to explore an issue and dig out what is key in it. The Metaplan process can be described in five steps: 1) agreement on principles; 2) exploratory discussions; 3) development of the "dramaturgy"; 4) meeting / workshop; 5) post meeting follow-up.

The advantages of this method are related to the difficulties to: 1) reach common points of view and take actions to support participants' convictions; 2) lead the group into joint actions; and 3) discuss and clarify the objectives within the allotted time.

The disadvantages are related to: 1) the necessity of having a skilled leader; 2) the possible existence of some overlap of ideas due to unclear wording or inadequate group discussion; and 3) "Knowledgeable" individuals selected to participate may not represent all community subgroups.

Surveys

Survey is a widely used method to gather scientific information about how people feel about a particular issue [67]. They are useful in determining correlations between sets of beliefs and perhaps giving hints to cause and effect relationships.

The survey is based on information collected from a sample of the total community population. On the other hand, a survey can be administered to all people in a community or organization to provide everyone with an equal opportunity to express themselves. The most commonly used survey methods are person-to-person interviews, drop-off and pick-up questionnaires, mail questionnaires, and telephone interviews. While each approach is somewhat different, the format is similar. Each asks an individual to supply attitudes, beliefs, behaviours, and attributes in response to specific questions.

The major advantages of this method are: it can be inexpensive, especially if volunteers are available to conduct the survey, or if data records about the particular issue already exist and can be reused; and a statistical sample can provide accurate information about a population.

The disadvantages of surveys are related to the: 1) assurance of statistical meaning, random samples must be carefully selected; 2) results may not be valid if surveys are not designed correctly; and 3) may require time and expertise to develop the survey, train interviewers, conduct interviews, and analyse results.

Delphi

Delphi is a method for structuring a group communication process so that it is effective in allowing a group of individuals, as a whole, to deal with a complex problem [68]. It allows free discussion of views without the influence of personal status; allows the iteration of personal views with controlled feedback and involves groups of people interested in a specific subject. This methodology is specialized in technology evaluation and can be used either for qualitative as well as for quantitative methods. The Delphi process can be described in three steps:

1. Selection of the “expert panel”: An expert may be considered to be an individual who has recognized expertise in a particular subject or may be anyone who can provide a worthwhile opinion on the subject in question. Thus, an expert panel is a group of people with expertise in some specific areas of the study able to give input to the formulation of the questionnaires in the next step;

2. The Delphi rounds: This step is composed by three or more rounds, depending on the situation:

    a. Round 1: In “classical” Delphi, the first round is completely unstructured, asking members to express any opinions about the current topic. The first round contains a synopsis of the issue in question together with the source and validity of the information upon which it is based;

    b. Round 2: With the results of the first round a questionnaire is constructed containing a series of statements or questions that respondents are invited to express an opinion on;

    c. Round 3: In the round 3, participants re-rank their initial statements in the light of the results of round two. Their own answers from the second round are feedback to the panellists so they may see their own answers in the light of the group’s overall response;

3. Results and conclusion: The results of the process are disseminated to the group as a well-researched guideline to best practice. This may then be used as a benchmark document for subsequent audit.

The Delphi method has the advantage of allowing participants to remain anonymous; it is free of social pressure, personality influence and individual dominance; and the questionnaires will be posted and answered using Internet, making the process inexpensive.

The disadvantages of this method are: 1) more time-consuming than the group process method; 2) should not be seen as a total solution to forecasting; 3) requires skills in written communication; and 4) requires adequate time and participant commitment.

2.1.3 **Hybrid Knowledge Elicitation**

Hybrid knowledge elicitation is an approach that aggregates both explicit and tacit knowledge elicitation characteristics. Tagging process is an example of hybrid knowledge elicitation, since users in such process are able to provide their own view, like in a kind of tacit knowledge elicitation. By other side, since tagging is able to aggregate several views in a formalised structure, it enables the possibility of extracting statistical patterns from such knowledge. This accomplishes the other elicitation kind, the explicit knowledge elicitation. To clarify what tagging is and for what it is used, further clarifications are presented.

Nowadays, the World Wide Web (WWW) has advanced to what is called Web2.0. The term Web 2.0 is associated with web applications that facilitate participatory information sharing, interoperability, user-centred design [69], and collaboration in the WWW [70]. By other words, Web 2.0 is a concept that shapes the online landscape into a platform for social connection, sharing, dialogue, and multimedia enhanced interaction. The goal is to promote social networking and creativity [71]. Inside these characteristics appeared the tagging process.

Tagging is a categorisation process that enables users to categorise resources on the web. It could be said that it also represents a kind of offloading intellect onto the web. There are three main entities that compose any tagging system: the users of the system (people who actually do the tagging); the tags themselves; and the resources being tagged. Each of these can be seen as forming separate spaces consisting of sets of nodes, which are linked together by edges (Figure 2.1) [72].

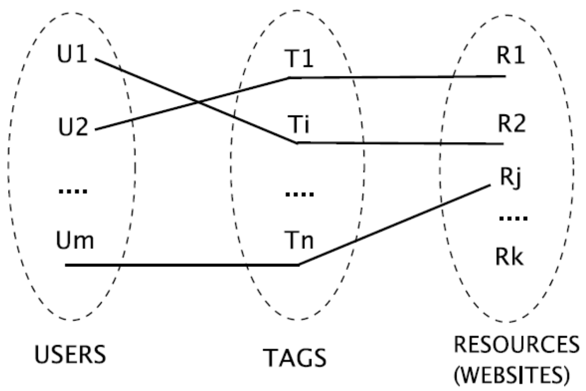


Figure 2.1 – Tripartite graph structure of a tagging system. An edge linking a user, a tag and a resource (website) represents one tagging instance [72]

Thus, tagging services allow a participant to associate freely determined keywords (called tags) with a particular resource [73]. Tags are simple pieces of data (terms or neologisms in natural language) that describe the item or resource tagged. Tags provide information about the item or resource as well as make it easier to see related items, i.e., that have the same tag [74].

Tagging is considered a categorization process, in contrast to a pre-optimized classification process used in ontologies. Jacob in [75] defines the distinction between categorization and classification in the following way: "Categorization divides the world of experience into groups or categories whose members share some perceptible similarity within a given context. This context may vary and with it the composition of the category is the very basis for both the flexibility and the power of cognitive categorization" while "classification involves the orderly and systematic assignment of each entity to one and only one class within a system of mutually exclusive and non-overlapping classes. It mandates consistent application of these principles within the framework of a prescribed ordering of reality [75]". Tagging systems allow much greater malleability and adaptability in organizing information than do formal classification systems [73].

In additional, there is the involuntary collaborative tagging, which relies on human knowledge, as opposed to an algorithm, to directly connect terms to documents before a search begins, and so relies on the collective intelligence of its human users to pre-filter the search results for relevancy. Folksonomies is what result from this tagging process. It represents the relations network of the various spaces of the tagging process. Folksonomy concept is also described later under "Knowledge Representation" section (chapter 2.2.1).

Figure 2.2 – Tag cloud example

Tag clouds are a result of such folksonomies. It is a special view of such gathered information. Tag Clouds are not only a great product of folksonomy classification [71]. Tag clouds are also, simple visualizations of term frequencies. A tag cloud allows you to see common terms in a text by grouping like terms together and emphasizing frequent terms [76]. Figure 2.2 presents an example of a tag cloud from the text written in this subchapter (2.1). They can also serve as a 'gateway' or 'doorway' into a section of the related web site or other collective content related to a specific term. Some people even use them as a sitemap alternative (each word or phrase in the tag cloud is then a link) [71].

2.1.4 **KA Conclusions**

The main point recorded from this research topic is that human intervention is always needed when the objective is to elicit knowledge from tacit or explicit format to a formalised solution as an ontology. However, even needing some human control, there are good solutions to extract knowledge from text to an ontology. Concerning the tacit knowledge elicitation, there are some appropriated methods, as

QICM that could contribute to organize knowledge engineers discussions and facilitate knowledge gathering. Such methods could be very useful for semantics interoperability resolution, especially when it is needed to establish agreements on concepts definitions between several actors as enterprises.

In additional, it is noticeable that there is the tagging process. Such process could act as a kind of hybrid knowledge elicitation solution because it uses both above solutions in its process of acquiring knowledge from users. When users perform the tagging process they are supplying information to the systems (web). Later, since such information provided is stored in a specific structure, it facilitates its further use in specific reasoning solutions able to extract patterns. Such acquired patterns represent new knowledge that came initially from humans.

Tagging data offers an interesting window into the intersection of formal reasoning and semi structured data with context-dependent semantics [77]. This promises several benefits and an interesting phenomenon since semantics can emerge from all the participants in a kind of a common view.

2.2 Knowledge Modelling (KMo)

Knowledge Modelling (KMo) packages combinations of data or information into a reusable format for the purpose of preserving, improving, sharing, aggregating and processing knowledge to simulate intelligence [78].

Consequently, this section starts by presenting some KR structures to be used for storing knowledge in databases or files, able to use any kind of KR language to facilitate further knowledge use and interpretation. Then, it is presented a set of ontology building methodologies and tools available to perform such KMo process. And finally, it is introduced a set of collaborative ontology building methodologies, with the purpose to present what are the existent solutions or approaches to perform a multi actor ontology building process.

2.2.1 **Knowledge Representation**

KR studies the formalisation of knowledge and its processing within machines. Techniques of automated reasoning allow a computer system to draw conclusions from knowledge represented in a machine-interpretable form [79].

A Knowledge Representation Element (KRE) is an element that facilitates the formal representation of the knowledge in a specific domain. In the following it is presented some examples: dictionary, glossary, taxonomy, thesaurus, ontology, and KB. The diagram of the Figure 2.3 illustrates the KREs relations existent in the two identified paths to build knowledge of a specific domain. The Figure 2.3 a) represents the taxis path and, the Figure 2.3 b) the folk's path to build knowledge. These nomenclature are explained by the own definition of the terms used: "taxis" comes from the Greek and means arrangement and to place in order [80], which in this specific case represents a specific arrangement, structured, of the lexicon to build knowledge; folks means common people of a society [81], which in this specific case represents the knowledge building by common people in their traditional way of understanding and characterising things.

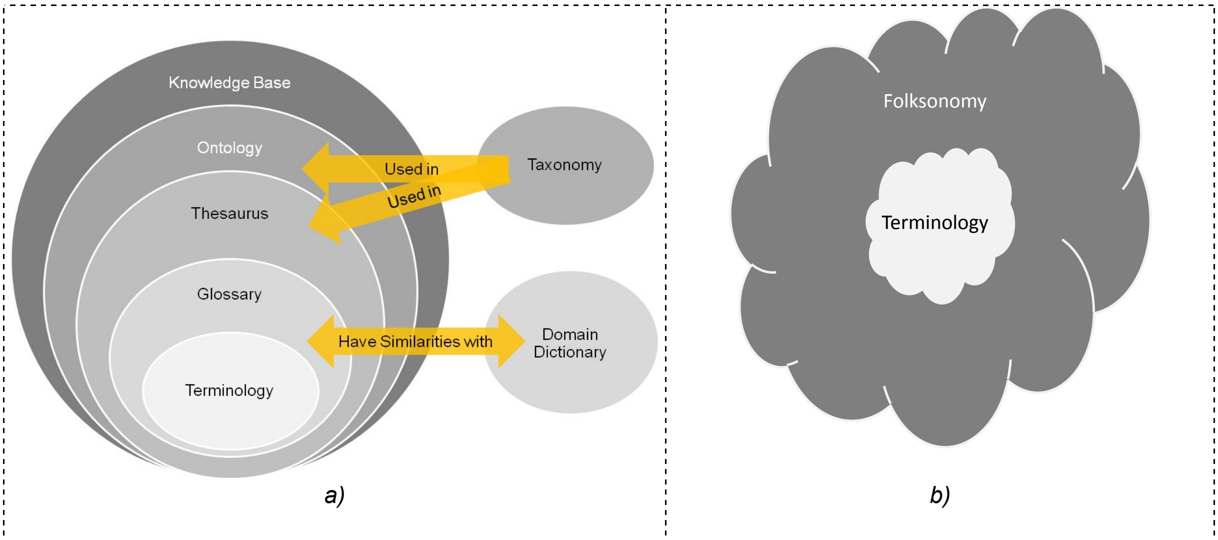


Figure 2.3 – KREs Relations in the: a) taxis; and b) folks; path to build knowledge

In linguistics, the lexicon of a language is its vocabulary, including its words and expressions [82]. For a human, knowing a language implies having a mental lexicon, i.e. a memorized set of associations among sound sequences, their meanings, and their syntactic privileges [83]. The Lexicon Settlement represents a domain knowledge acquisition, which comparatively to the human language apprentice phase could be represented in computer science as a semantic organized structure with definitions. A KB composed by ontologies and thesaurus structured by a similar taxonomy that aggregates a glossary of terms in a specific domain can represent such semantic organized structure establishing in this way the lexicon of a specific domain (Figure 2.3) [84].

A concept is a cognitive unit of meaning—an abstract idea or a mental symbol sometimes defined as a "unit of knowledge," built from other units, which act as a concept's characteristics. A concept is typically associated with a corresponding representation in a language or symbology such as a single meaning of a term. Conceptualisation is the process related to the generation of units of knowledge. Is in the pursue of building several “units of knowledge” that KR concept works, accomplishing at the end of its process a KB.

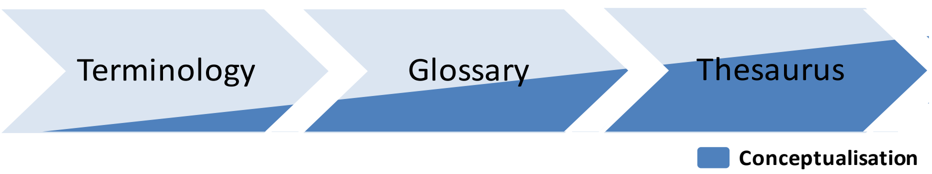


Figure 2.4 – Level of conceptualisation presence in KREs

In order to have a KB it is needed to gather the knowledge in an explicit way composed by terminology, glossary, thesaurus and ontologies supported by taxonomies, domain dictionaries and human tacit knowledge. Such KREs present a distinct level of conceptualisation. Figure 2.3 represents it accordingly to the level of conceptualisation that each one has. There is an increase of its presence from terminology to the KB (Figure 2.4 represents an illustration of this presence).

Terminology

Terminology is the study of terms and their use. Terms are words and compound words that are used in specific contexts. Terminology therefore denotes a more formal discipline which systematically studies the labelling or designating of concepts particular to one or more subject fields or domains of human activity, through research and analysis of terms in context, for the purpose of documenting and promoting correct usage [85].

Domain Dictionary

A dictionary is a book of alphabetically listed words in a specific language, with definitions, etymologies, pronunciations, and other information; or a book of alphabetically listed words in one language with their equivalents in another, also known as a lexicon.

Sofa: *Comfortable seat with raised arms and back filled or covered with soft material and long enough for two or more people to sit on.*
Related words: *Couch*
Translation: Fr. *Canapé*; Port.Eu *Sofá*; Esp. *Sofá*

Figure 2.5 – Extract of a domain dictionary

A domain dictionary has been found to be one of the most useful tools for a domain analysis. The dictionary lessens a great deal of miscommunication by providing users with information: 1) in a central location to look for terms and abbreviations that are completely new; 2) where definitions of terms are used differently or in a very specific way within the domain [86][22]. Figure 2.5 illustrates an example of an extract of a domain dictionary.

Glossary

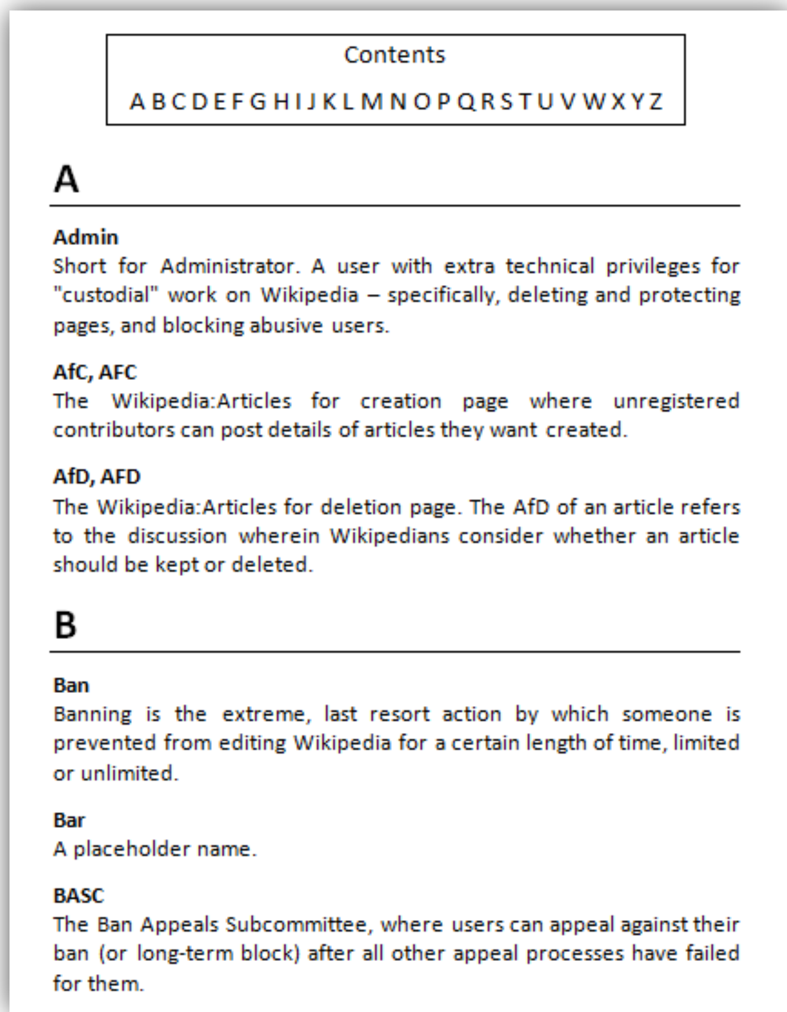
Glossary is a list of specialized terms, mostly in alphabetic order, that sometimes are unique to a specific subject. Each term is composed by its corresponding description. It includes descriptive comments and explanatory notes, such as definitions, synonyms, references, etc.

A glossary can be used when communicating information in order to unify knowledge sharing. A glossary is understood as a set of terms and their definitions, and is bound to the source document or domain where these definitions are set (Figure 2.6).

Taxonomy

It has become fashionable in certain circles to apply the term in a wider, more general sense, where it may refer to a classification of things or concepts, as well as to the principles underlying such a classification [87].

A taxonomy is a classification system that categorizes all the information in a class/subclass relationship, representing a simple tree structure. At the top of this structure is a single classification, the root node that applies to all objects. The root node represents most general category of all things that the domain is related to. Nodes below this root are more specific classifications that apply to subsets of the total set of classified objects [87]. For any category, each subcategory is a taxonomy. Each child is a subset of the parent. The intersection of each pair of children, in same level, is empty. Any path from the root to a leaf is called a branch. Since taxonomies are transitive, child nodes aggregates parents categories by heritage, e.g. tiger and cat can be a child class of the mammals' class, thus they are mammals also [88].



Contents

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

A

Admin
Short for Administrator. A user with extra technical privileges for "custodial" work on Wikipedia – specifically, deleting and protecting pages, and blocking abusive users.

AfC, AFC
The Wikipedia:Articles for creation page where unregistered contributors can post details of articles they want created.

AfD, AFD
The Wikipedia:Articles for deletion page. The AfD of an article refers to the discussion wherein Wikipedians consider whether an article should be kept or deleted.

B

Ban
Banning is the extreme, last resort action by which someone is prevented from editing Wikipedia for a certain length of time, limited or unlimited.

Bar
A placeholder name.

BASC
The Ban Appeals Subcommittee, where users can appeal against their ban (or long-term block) after all other appeal processes have failed for them.

Figure 2.6 – Example of a Glossary

Figure 2.7 illustrates an example of a taxonomy creation. It starts by the gathering of the terms in a domain (Figure 2.7 a)), then the terms are organised beginning from the more generic one(s) or by one that could aggregate the others through a relation on the meaning. This is done iteratively until have all the terms associated to the overall tree structure (Figure 2.7 b)).

Thesaurus

A thesaurus is a structure that manages the complexities of terminology and provides conceptual relationships, ideally through an embedded classification. A classification is a structure that organizes concepts into a hierarchy, possibly in a scheme of facets. A monolingual thesaurus has terms from one language, a multilingual thesaurus from two or more languages [89].

The thesaurus can be then represented by a set of classes representing domain reference concepts with associated meanings about a domain in a semantic related structure. A thesaurus is like a taxonomy of domain concepts composed by its reference meanings.

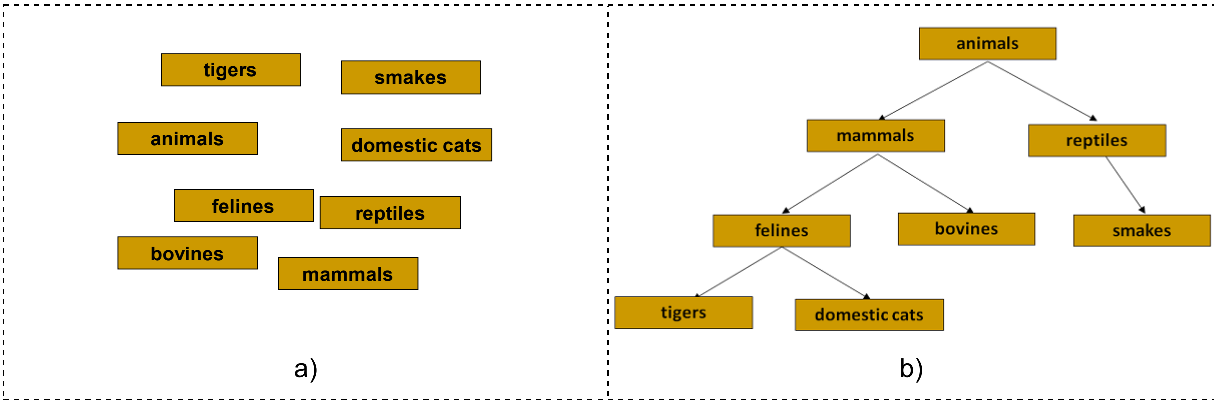


Figure 2.7 – Example of a Taxonomy Construction

Ontology

Gruber stated in 1993 one of the most well known description about what is an ontology: An ontology is a formal, explicit specification of a shared conceptualisation [90]. The term is borrowed from philosophy, where ontology is a systematic account of Existence [90]. Where, a 'conceptualisation' refers to an abstract model of some phenomenon in the world, which identifies the relevant concepts of that phenomenon. 'Explicit' means that the type of concepts used and the constraints on their use are explicitly defined. And, 'Formal' refers to the fact that the ontology should be machine understandable. It can be translated into some form of logic, usually first order logic. For example, the statement “Every regular chair has 4 legs.” can be expressed in logic as “for all x, if x is a regular chair then x has 4 legs.” And finally, 'Shared' reflects the notion that ontology captures consensual knowledge that is not restricted to the knowledge view of some individual, but reflects a more general view shared and accepted by a group.

Recently, it has been seen an explosion of interest in ontologies as artefacts to represent human knowledge and as critical components in knowledge management, the Semantic Web[11], business-to-business applications, and several other application areas. Various research communities commonly assume that ontologies are the appropriate modelling structure for representing knowledge. However, little discussion has occurred regarding the actual range of knowledge an ontology can successfully represent [91].

Taking the example of the taxonomy presented before, ontology is like the enrichment of such structure, with properties and rules that relates its concepts with the purpose of representing a segment of the reality following the view of a determined group.

Ontologies are now being recognised as important components of information systems and

[11] The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. It is a collaborative effort led by W3C with participation from a large number of researchers and industrial partners. It is based on the Resource Description Framework (RDF), which is a standard model for data interchange. Retrieved from the web at May 2012: http://www.w3.org/2001/sw/; and http://www.w3.org/RDF/.

information processing. In the areas of KR and reasoning and of conceptual modelling, it has long been recognised that conceptualising a domain is a prerequisite for understanding the domain and processing information about the domain, especially in the case of large, non-trivial domains [92].

In the context of the Semantic Web, ontology is an enabling technology - a layer of the enabling infrastructure - for information sharing and manipulation. The approach is simple: parties who have software/data/services to offer identify some common conceptualization of the data; they specify that conceptualization as clearly they can; they build systems that interoperate on those specifications. This is standard-issue information technology, with the twist that ontologies are specifications of the conceptualizations at a semantic level [22][93].

An ontology provides a vocabulary that describes a domain of interest and a specification of the meaning of terms used in the vocabulary. By defining shared and common domain theories, ontologies help both people and machines to communicate concisely, supporting the exchange of semantics and not only syntax [51]. In computer and information science, ontology is a technical term denoting an artefact that is designed for a purpose, which is to enable the modelling of knowledge about some domain, real or imagined. So, it represents a KR used to capture information and knowledge about a subject, generally within the structure of a semantic network, consisting of a diagram composed of nodes and arcs. It could be defined a class hierarchical ontology representing concepts, objects or entities characterized by their properties [94].

Knowledge Base

Ontologies aggregate semantics, rules and characteristics able to classify and represent a segment of the reality. An ontology is the conceptual system, which underlies a particular KB [95] By other words, ontologies provide some structure for development of KBs, serving to define its models [96].

Figure 2.8 – Real existent chair: POÄNG model produced by IKEA[12]

When ontologies are associated with real instances/individuals it becomes a KB. As an example, an ontology is able to represent any kind of chair, consequently if it starts to represent real chairs as the one presented in the Figure 2.8 it becomes to be in this case, a KB of chairs.

[12] Retrieved from the web at August 2011: http://www.ikea.com/us/en/catalog/products/S69805416

Folksonomy

Recently an emergent concept is growing in the Semantic Web called Folksonomies. It arises from data about how people associate terms with content that they generate, share, or consume. Folksonomy (from folk and taxonomy) is a neologism for a practice of collaborative categorization using freely chosen keywords [77]. The idea of a folksonomy is to allow the users to describe a set of shared objects with a set of keywords of their own choice. The interesting observation is that when users (folks) do their tagging in a public space, the collection of their keyword/value associations becomes a useful source of data in the aggregate, hence the term "folksonomy", the labelling of lots of things by people in a social context.

It is important to note that in terms of KR, the set of keywords used in such labelling process can only be considered as a list of terms. The set of rules is not fixed, in fact, the users form no explicit agreement at all about the use of words, and also important is that there is no one-to-one correspondence between concepts and keywords (terms used). Notice that it is not always possible for the users to express a complex concept with a single keyword and thus they may use more than one tag to express the concept association that the item brings up in them.

In contrast with taxonomies, that limit the dimensions along which one can make distinctions, and local choices at the leaves are constrained by global classifications in branches, folksonomies are massively dimensional and there is no global consistency imposed by current practice. Things are easy to tag, there is no wrong answer, and the emergent patterns give insight into collective attention [77].

This attack on "ontology" is really an attack on top down categorization as a way of finding and organizing information and the praise for folksonomy is really the observation that we now have an entirely new source of data for finding and organizing information: user participation. For the task of finding information, taxonomies are too rigid and purely text-based search is too weak. Tags introduce distributed human intelligence into the system [93].

2.2.2 Ontology Building Methodologies

There is no universal methodology to build an ontology and designers of ontologies themselves apply different views of the same domain during ontology development. This yields semantic heterogeneity at ontology level, which is one of main obstacles to semantic interoperability [97].

Academic and industrial communities have been developing methodologies to guide in the process of ontology building. Based on Gómez-Pérez et al. [98][99], the author clustered the following set of characteristics to clarify ontology building methodologies, which are: 1) starting from scratch (S); 2) reengineering (Re), i.e. ontology building based on existing ones; 3) cooperative building (Co), i.e. the actors should be able to participate in the process; 4) merge methods (Me), i.e. ontology merging, integration, use and mapping.

Ontology merging allows the development of a more general ontology about a subject, by gathering knowledge from several other ontologies in that same subject. Ontology integration reuses other ontologies, while, each integrated ontology is about a different domain either from the resulting ontology. Ontology use builds an application using one or more ontologies – such process does not have any resulting ontology [100]. Ontology mapping is an activity that attempts to relate the vocabulary of two ontologies that share the same domain of discourse [101].

Methodologies	S	Re	Co	Me	Tools	Languages
Cyc	✓				KAON	OCML
Uschold & King	✓				OILed	**OWL**
Grüninger & Fox	✓				OntoEdit	**DAML+OIL**
KACTUS	✓				WebODE	OIL
METHONTOLOGY	✓	✓			**Protégé**	**RDF(S)**
SENSUS	✓	✓			Ontolingua	**XML**
CO4			✓		Server	KIF
(KA)²			✓		Ontosaurus	Ontolingua
On-To-Knowledge	✓				WebOnto	**OKBC**
FCA-merge				✓	VOID	LOOM
PROMPT				✓	Apollo	**FLogic**
ONIONS				✓	CODE4	
UPON	✓				NeOn	
NeOn	✓	✓	✓	✓	Toolkit	

Legend: ⟶ method. direct relation; —— languages used by the tool
S – Scratch; Re – Reengineering; Co – Cooperative; Me – Merge

Table 2.2 – Ontology Building Methodologies Analysis

In the Table 2.2, it is presented a set of available ontology building methodologies, categorized by the identified ontology building characteristics. It presents also in the middle, a set of tools used for ontology building, where some of them are related to the presented methodologies (relation represented by an arrow). Then in the right part, it is presented a set of languages, the ones in bold are used by the Protégé tool. Protégé tool is also in bold due to its relevance to this research. NeOn methodology is as well in bold, due to be the only one that has all the mentioned ontology building characteristics.

The CYC methodology consists on extracting, by hand, common sense knowledge that is implicit in different sources. Once enough knowledge in the ontology is available, new common sense knowledge can be acquired either using natural language or machine learning tools [102].

The Uschold & King proposed some general steps to develop ontologies, which are: to identify the purpose; to capture the concepts and the relationships among these concepts and the terms used to denote both of them; to evaluate and to document the ontology [103].

Grüninger & Fox methodology is based on identifying the main scenarios and the competency questions, followed by extracting relevant concepts and relations and formalizing it [104].

In the KACTUS project an ontology methodology was proposed, which its building process is based on a KB application, by means of an abstraction process [105]. This methodology has VOID as supporting tool.

Methontology is a methodology for building ontologies either from scratch, or by a re-engineering process. The Methontology framework enables the construction of ontologies at the knowledge level. It includes: identification of the main ontology development process activities (i.e., evaluation, configuration management, conceptualization, integration, implementation, etc.) [106]. This methodology was accomplished with a tool called WebODE.

Sensus is a methodology, which follows a top-down approach for deriving domain specific ontologies from huge ones. The process is based on a first manual identifying of terms that are relevant to a particular domain and a consequently automatic extraction of the related hierarchically structured of the previous identified terms. The result is a skeletal foundation for a KB [107].

The CO4 methodology foundation is based on a protocol defined from identified problems concerning collaborative construction of ontologies. The goal is to reach consensus among several KBs and it is based on the main idea that people can discuss and commit about the knowledge introduced in the KBs. These KBs are built to be shared, and they have consensual knowledge, hence they can be considered ontologies [108].

The (KA)2 methodology is based on ontologies development in a joint effort by a group of people at different locations using the same templates and language. The process is based on ontology coordinating agents which distribute a template among the ontology topic agents (experts in different topics). Once the ontology coordinating agents got all the portions of the ontologies from the ontology topic agents, they integrate them, activity that benefits from the presence of a common pattern [109].

On-To-Knowledge methodology consists on generating editable ontologies automatically from natural language documents followed by human interaction in order to ensure quality of the results [106]. This methodology is accomplished by the tool OntoEdit.

The method of FCA-Merge is guided by application-specific instances of the given source ontologies that are to be merged. Natural language processing and formal concept analysis techniques are applied, in order to derive a lattice of concepts. The generated result is then explored and transformed into the merged ontology with human interaction [110].

PROMPT is an algorithm that provides a semi-automatic approach to ontology merging and alignment. PROMPT performs some tasks automatically and guides the user in performing other tasks for which his intervention is required [111]. PROMPT is a plug-in of the tool Protégé.

ONIONS is a methodology for conceptual analysis and ontological integration or merging of terminologies. ONIONS aims to provide extensive axiomatization, clear semantics, and ontological depth in the domain terminologies that are to be integrated or merged [112].

UPON is a Unified Process for Ontology Building based on workflows of iterations related to the ontology building requirements, analysis, design, implementation and testing [113].

NeOn is a project involving 14 European partners and co-funded by the European Commission’s Sixth Framework Programme under grant number IST-2005-027595. NeOn had the aim to advance the state of the art in using ontologies for large-scale semantic applications in the distributed organizations. Particularly, it had the main objective of improving the capability to handle multiple networked ontologies that exist in a particular context, are created collaboratively, and might be highly dynamic and constantly evolving [114]. From such project it was defined the NeOn methodology that, as already stated, aggregates all the ontology building characteristics. NeOn identified a set of different building ontology scenarios accomplished with documented procedures to guide users developing their ontologies. NeOn produced a tool to support its methodology, the NeOn Toolkit.

In general, there is no correspondence between ontology building methodologies and tools for ontology management, as could be observed on Table 2.2. Although there is a few that has a direct tool which implements its related methodology process. Thus, are these ones that were considered on this literature review: OntoEdit; WebODE; VOID; Protégé and NeOn. OntoEdit is the old tool that On-To-Knowledge was related to, now its successor is OntoStudio [115], which is a shareware tool. Since NeOn Toolkit is the sister project of the ontology engineering environment OntoStudio developed by ontoprise [116], which is shareware, author only performed a comparison between NeOn and Protégé.

The NeOn Toolkit offers most of the commercially available plug-ins and also provides access to many community-built plug-ins via the update-site mechanism. The toolkit is an ontology engineering environment for modelling ontologies. It is based on Eclipse[13] and a modular design that is extensible by plug-ins contributed by external developers. The NeOn Toolkit supports two different ontology languages: F-logic and (Web Ontology Language) OWL [116]. In additional NeOn Toolkit is managed and distributed by the not-for-profit organization NeOn Technologies Foundation.

Protégé is a free, open-source platform that provides a growing user community with a suite of tools to construct domain models and knowledge-based applications with ontologies. At its core, Protégé implements a rich set of knowledge-modelling structures and actions that support the creation, visualization, and manipulation of ontologies in various representation formats [117]. As it is emphasized in the right column of the Table 2.2, Protégé works with a vast list of languages.

Author tested both (Protégé and NeOn) tools in developing a set of simple ontologies and he concluded that they offer almost the same functionalities and characteristics with a similar level of active communities. However, author pointed out some characteristics, which both differ from each other, and which he thinks are relevant for developing ontology-based prototypes. Such characteristics are related to: plug-ins availability and installation easiness, documentation availability and usability, maturity of the tool, quantity of import export languages.

In this list of characteristics, NeOn only surpasses Protégé in one that is documentation availability. Documentation offered by NeOn, perhaps due to being a more recent tool, it is more detailed and better described. Concerning plug-ins both have a big list of them, but, installations procedures of NeOn are more difficult than in Protégé. In one specific case, related to the ontologies server plug-in, NeOn indicates one which is shareware, which belongs to OntoStudio. Then, Protégé is an older tool, which is able to offer more mature support at least concerning to the how an open-source community work. By the last, Protégé is able to import and export a higher number of different ontologies languages. However, in a kind of positive point to both communities, it is interesting to confirm ontologies interoperability ability between both, when used ontology languages standards (e.g. OWL).

[13] Eclipse is an open source community, whose projects are focused on building an open development platform comprised of extensible frameworks, tools and runtimes for building, deploying and managing software across the lifecycle. Retrieved from the web at May 2012: http://www.eclipse.org/org/.

2.2.3 **Collaborative Ontology Building Methodologies**

There are many collaborative ontology building tools and methodologies proposed. Those identified to have characteristics able to facilitate collaborative human knowledge acquisition (i.e., tacit knowledge) into explicit KR are herein described.

The Iterative Collaborative Ontology Construction (ICOC) scheme [118], supports the online collaborative knowledge contribution. It uses a wiki-like application that allows users to collaboratively integrate their knowledge to build a new ontology. A Delphi-like method is then applied to converge the answers to an automatic generated questionnaire to construct a uniform ontology. This scheme will be processed iteratively until all relations are converged.

Holsapple et al. in “A Collaborative Approach to Ontology Design” [119], used a Delphi-like method to structure collaboration in the direction of consensus. The method used is composed by four phases: 1) in the Preparation phase the idea is to define design criteria, determine boundary conditions, and determines evaluation standards; 2) in the Anchoring phase is specified the initial ontology that will seed the collaborative effort; 3) the Iterative Improvement phase pretends to identify diverse panel of participants that provide their critiques and comments on the ontology to a leader who revises the ontology addressing the feedback from the users. This process iterates until a consensus is reached; 4) in the Application phase the idea is to demonstrate the uses of the ontology.

Collaborative Protégé [120] is an extension of the existing Protégé tool to support collaborative ontology editing. The traditional Protégé tool is used to create and edit ontologies but without collaboration between users. With Collaborative Protégé users are able to join a project through a server, and if they have permission, they can edit ontologies through desktop or web Protégé clients. All the changes made by a user will be immediately reflected on the repository, thus online users will see the changes. Additionally, Collaborative Protégé allows other operations such as track changes, chat, leave comments on ontology changes, voting, etc.

The OntoWiki [79] tool is another kind of collaborative environment that supports agile Knowledge Engineering. To reach social collaboration characteristics, OntoWiki uses a particular set of functionalities: Change tracking, Commenting, Rating, Popularity and Activity/Provenance. Additionally, OntoWiki facilitates the visual presentation of a KB as an information map, with different views on instance data. It enables intuitive authoring of semantic content, with an inline-editing mode for editing RDF content, similar to WYSIWYG (What You See Is What You Get) for text documents [121].

By the last, in the INTEROP research project [122], it was used another collaborative approach to build a glossary. Since, as presented before, glossary is a KRE needed to be defined on the process to build an ontology, which by consequence justifies the presence of this approach here. It starts by using the OntoLearn TermExtractor module to extract terms from specific INTEROP related documents. Then, the list of terms is reviewed by a set of domain experts to refine the glossary. On that process, the terms could be reviewed, rejected, accepted or ignored by the reviewers in an iterative process until the majority of voters approve all definitions.

2.2.4 KMo Conclusions

Since semantic interoperability resolution involves several actors as enterprises trying to reach a common view in a domain, require methodologies focused in plural actors. Consequently, the main conclusion from this research topic is related to the need of having an appropriate methodology able to support KMo supported by the use of existent technologies for ontology collaborative building and possibly enriched with QICM. In conclusion, such methodology should be able to facilitate the ontology building from scratch, and by reengineering of existent ones. It should also facilitate a cooperative building of ontologies using, when needed, merging or integration functionalities supported by the establishment of mappings between existent ontologies and the new built ones, enabling the possibility of keeping in use old semantic representations.

It was identified that NeOn methodology aggregates all these mentioned characteristics, however, is the Protégé tool that facilitates more the collaborative ontology building through its available plug-ins. Moreover, NeOn do not aggregates all the mentioned characteristics in a methodology as an all, supporting enterprises to establish a common view on their semantics in a way, which establishes mappings tables at the same time that semantic agreements are reached. Thus, its reuse in communications between old semantic and new semantic systems is not facilitated. From these, it is identified the need of having a methodology focused on supporting the building of a reference ontology able to represent all the actors semantics, but at the same time that could let enterprises still work with their traditional and old semantics.

Finally, this research work helped to identify as stated by Halpin et al. [72], "It seems quite plausible that folksonomies and ontologies, which are merely new incarnations of the age-old distinction between categorization and classification respectively, are not mortal enemies, but fundamentally compatible, as tagging-based categorization in our data exhibits emergent consensus". Thus both KREs (ontologies and folksonomies), whose enable the taxis or folks path to build knowledge, have their own specific usability and purpose however they can be used in a compatible way to improve the knowledge acquire process.

2.3 Knowledge Use (KU)

This stage is centralised **on the use of the stored knowledge to facilitate interoperability of the systems**. Thus the main objective is potentiating semantic adaptability of the system by using the knowledge about the related domain.

Due to the worldwide diversity of communities, a high number of ontologies have appeared representing the same segment of reality. Thus, various parties of a same domain often do not understand each other because of the use of different ontologies. To force such parties to adopt the same ontology, even if it is based on standards (e.g. AISI, DIN or ISO)[14], does not work in most of the

[14]AISI – AISI the American Iron and Steel Institute has established standards for steel compositions. - http://www.steel.org
DIN - Deutsches Institut für Normung e.V. (DIN; in English, the German Institute for Standardization) is the German national organization for standardization and is that country's ISO member body. - http://www.din.de

cases [1]. An ideal solution would be to keep the terminology and classification in use by each one, and to use a 'reference ontology' to be the intermediary in the communications between them. Borgo et al. [123] characterised 'reference ontologies' (more recently, foundational ontologies) as rich, axiomatic theories whose focus is to clarify the intended meanings of terms used in specific domains [124]. Additionally, the introduction of this reference ontology would enrich the community where each party should feel motivated to be part of the group, since they have the possibility of keeping their own definitions, i.e. their own ontology, though they have to contribute to the reference ontology building process. This establishes the knowledge alignment between each party and the reference ontology that provides a consensual common vocabulary for describing the knowledge in the domain [125][126].

Even so, new organizations are constantly entering and leaving such collaboration networks, leading to a constant fluctuation and evolution of system models. All these factors are making interoperability difficult to sustain [127]. Due to this constant knowledge changes, ontologies and model mappings are not static and there is always some information to add to a KR system. An ontology has to manage its dynamics, being able to learn during its existence and be in a constant update [29]. Thus, data, semantic, and structural mappings could be modelled as traceable tuples and integrated in KBs dedicated to managing mismatches. Next sub-sections summarize different ways to represent and formalize Model Morphisms (MoMo) and semantic matching, and then it is defined the concept of tuple for semantic and structural mapping.

2.3.1 Information Models and Associated Concepts Models

Either being used in the form of traditional databases, architectural models, or domain ontologies, information models can be described on multiple formats, languages, expressiveness levels, and for different purposes [125][128][129]. A model can be characterized according to four dimensions [130]: Metamodel - the modelling primitives of the language for modelling (e.g. OWL, XSD[15]) are represented by a set of labels defined in the metamodel; Structure - corresponding to the topology associated to the model schema; Terminology - the labels of the model elements that don't refer to modelling primitives; Semantics - given a "Universe of Discourse", the interpretations that can be associated with the model. This way, model operations can be classified as acting on any of these dimensions.

MoMo	Formalization	Classification
Mapping: $\theta(A,B)$	$\forall\ A,B \in M: \theta(A,B) \subseteq Sub(A) \times Sub(B)$	Non-altering
Transformation: $\tau: A \times \theta\ \rightarrow B$	$\forall\ A,B \in M: if\ \exists\theta(A,B) then\ \tau(A,\theta) = B$	Model altering

Table 2.3 – Cases of Model Morphisms

Model Morphisms

In mathematics, "Morphism" is an abstraction of a structure-preserving map between two mathematical structures. It can be seen as a function in set theory, or the connection between domain and co-domain in category theory [130]. Recently, this concept as been gaining momentum applied to computer science, namely to systems interoperability. This new usage of "morphism" specifies the relations (e.g. mapping, merging, transformation, etc.) between two or more information model

ISO - International Organization for Standardization – www.iso.org
[15] XSD - XML Schema Definition

specifications (M as the set of models). Therefore, a MoMo describes a model operation.

In this context, the research community identifies two core classes of MoMo: non-altering and model altering morphisms [130][131]. As evidenced in Table 2.3, in the non-altering morphisms, given two models (source *A* and target *B*), a mapping is created relating each element of the source with a correspondent element in the target, leaving both models intact. In model altering morphisms, the source model is transformed using a function that applies a mapping to the source model and outputs the target model [132]. Other relations, such as the merge operation, can also be classified as model altering morphisms.

Mismatch		Description
Lossless	Naming	Different labels for same concept
	Granularity	Same information decomposed (sub) attributes (see Figure 2.10)
	Structuring	Different design structures for same information (see Figure 2.9)
	SubClass-Attribute	An attribute, with a predefined value set (e.g. enumeration) represented by a subclass hierarchy
	Schema-Instance	An attribute value in one model can be a part of the other's model schema (see Figure 2.9)
	Encoding	Different formats of data or units of measure (e.g. USD and EUR)
Lossy	Content	Different content denoted by the same concept
	Coverage	Absence of information
	Precision	Accuracy of information (see Figure 2.9)
	Abstraction	Level of specialisation (e.g. "Car" and "Ford")

Table 2.4 – Semantic Mismatches (based on [133])

To respond to the constant knowledge and model changes on heterogeneous and dynamic networks, it is required to use a more detailed and traceable mapping format that provides a semantic "link" between two different models and its components.

Semantic Mismatches

Mismatches are inconsistencies of information that result from "imperfect" mappings. Due to the differences among models referred before, almost in every case, a MoMo leads to a semantic mismatch, which can either be lossy or lossless depending on the nature of the related model elements (Table 2.4): In lossless cases, the relating element can fully capture the semantics of the related; while in lossy mismatches a semantic preserving mapping to the reference model cannot be built [133].

Figure 2.9 – Mismatch examples

This notion of mismatch can bring a semantic meaning to the type of the relationship being established in the mapping. However, the envisaged semantic "link" between two different models needs to account for more than inference of a meaning. It needs to be represented through a formal expression that is traceable and parseable by an intelligent system that can deduce and recommend mapping readjustments, which might even change the mismatch type.

MoMo Formalisms

Model Morphisms, as envisaged before, are intended to introduce a method of describing relationships/transformations among models. Originally graph theory has been used, but other and theories can be considered to achieve the envisaged goals:

Classical Mathematics: Graph & Set Theory

Graphs are a common way to graphically present models, where the nodes are considered as a domain entity and the edges as relations between them. For the purposes of MoMo, model operations such as the ones of Table 2.4 can be described using a 6-tuple labelled oriented multigraph *(LDMGraph)* of the form $G=(V,E,s,t,lv,le)$, where: V is the vertex set of G; E is the edge set of G; $s:E \rightarrow V$, is a function that associates an edge with its source vertex; $t: E \rightarrow V$, is a function that associates an edge with its target vertex; $lv: V \rightarrow \sum V$, is a function that associates a vertex with its label; $le: E \rightarrow \sum E$, is a function that associates an edge with its label [130], [134]. This abstract view of models allows formal reasoning on their properties and on the properties of the model operations needed for their effective management.

As graphs, also sets can be used to represent models and operations using first-order logic, algebra and axioms. Being defined as a collection "*M*" of distinct objects "*m*", a set can represent objects, numbers, other sets, etc. [135]. Operations such as membership "$M1 \subseteq M2$", power "$P(M)$", union"$M1 \cup M2$", intersection "$M1 \cap M2$", complement "$M1 \setminus M2$", or cartesian product "$M1 \times M2$" are already well defined.

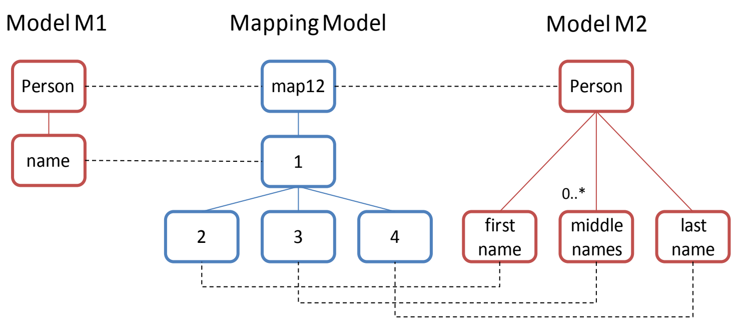


Figure 2.10 – Mapping as a model (map12)

<u>Mapping as a model: Model Management [136]</u>

This theory defends that a mapping between models *M1* and *M2* should be a model "*map12*" and two morphisms (one between "*map12*" and *M1* and another between "*map12*" and *M2*). Thus, each object "*m*" in the mapping can relate a set of objects in *M1* to a set of objects in *M2*. In this approach, instead of representing a mapping as a pair of objects, a mapping is represented as a set of objects (see Figure 2.10). Using concepts from classical mathematics, this approach enables to define complex algebra to describe major model operations such as match, compose, diff, model gen, or merge.

Mapping as a complex tuple: Matching [137]

The match operator takes two graph-like structures and produces a mapping between the nodes of the graphs that correspond semantically to each other. Mappings between these elements can be described using set-theoretic semantic relations instead of using traditional numeric coefficients. The meaning of concepts (not labels) within a model can determine equivalence "=", more "⊇" and less "⊆" general, as well as disjointness "⊥" relationships. Having this, a mapping element can be defined as a 4 level tuple <*IDij, ai, bj, R*> where: *IDij* is a unique identifier of the given mapping element; *ai* is the *i-th* node (or vertex) of the first tree; *bj* is the *j-th* node of the second tree; and *R* specifies the semantic relation which may hold between them.

The above methodologies seem to be powerful in terms of expressiveness of the morphism. However others exist, such as the composition of complex operations based on a catalogue of primitive transformations [138]. However, this approach is more focused on model altering morphisms.

2.3.2 Modelling Morphisms to Enable Sustainable Interoperability

So far, a proven approach to deal with interoperability relies on the usage of dedicated knowledge models and international standards acting as information regulators among organizations. However, due a complexity of reasons many organizations are still focused on P2P relationships, where each one tends to use its own data format and business rules, and handles as many mappings as the number of business partners [127].

Either case, after interoperability is first established and all morphisms defined, the set of organizations within a network demonstrate a period of stability exchanging e-messages following the established mappings [127]. At this stage, the networks display symmetry [139]. However, that might not be sustainable for long if business requirements change. Organizations are managed by people that have different opinions and backgrounds based on several factors such as culture, professional experience, family, etc. They manage, work, and are themselves customers of different organizations, which in turn have different systems that are structured according to several information models implemented on multiple software platforms. Therefore, all this heterogeneity leads in most cases, the network to experience problems because if just one of the network members adapts to a new requirement, the harmony is broken, and the network begins experiencing interoperability failure.

To mitigate that, context awareness and traceable morphisms are demanded in support of intelligence and monitoring and decision support systems must be considered in the construction of a framework that implements sustainable interoperability in cooperation networks.

Knowledge Enriched Tuple for Mappings Representation

Observing all previously explained technologies and methodologies for managing morphisms, the author considers that there is no perfect solution that can provide all the desired goals at once. Some are ideal for structural issues, others for semantics providing good human traceability, while others are more formal and mathematical based. Therefore, author proposes the usage of a 5-tuple mapping expression (equation 1), reusing some of the concepts explained before that formalizes the morphism between two model elements (a and b) and is enriched with semantic information that enables fast human readability, where $\forall\, A, B\ \in M, \exists a \in A\ and\ \exists b \in B\colon\ if\ M\ is\ an\ LDMGraph\ then\ a \in V(A) and\ b \in V(B)$.

Mapping Tuple: <*ID,MElems,KMType,MatchClass,Exp*> (1)

- *ID* is the unique identifier of the mapping tuple and can be directly associated with the a's vertex number: $IDi.j\_x\colon 1 \le i\ \le |V(A)|\ and\ 1 \le j \le \left|V\big(Sub(B)\big)\right| and\ x \in \mathbb{N}$. The depth of the sub-graph detail used in the mapping is not limited, and *x* is a counter for multiple tuples associated with the same concept;

- *MElems* is the pair *(a,b)* that indicates the mapped elements. If the *ID* specifies a mapping at the n-th depth level of the graph, *a* should be at the same level, i.e. *a.ai (for i =1..n);*

- *KMType* stands for Knowledge Mapping Type, and can be classified as: "*Conceptual*" if mapping concepts and terms; "*Semantics*" if mapping model schemas; and "*InstantiableData*" if the mapping is specifying instantiation rules.
    - *KMType= {Conceptual,Semantics,InstantiableData};*

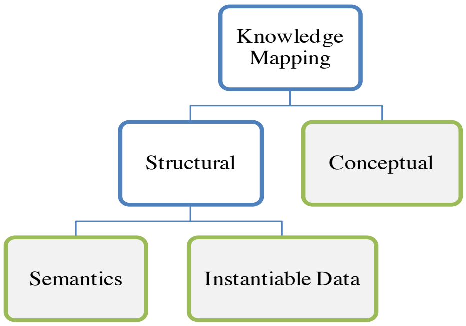


Figure 2.11 – KMType values

- *MatchClass* stands for Match/Mismatch Classification and depends on *KMType*, such as $\forall(a,b) \in MElems$:
    - $\forall KMType$, *if a=b,* the mapping is absolute and *MatchClass=Equal*;
    - if *KMType=Conceptual*, the mapping is relating terms/concepts, and $MatchClass \in \begin{Bmatrix} Equal, Naming, Coverage, \\ MoreGeneral, LessGeneral, Disjoint \end{Bmatrix}$ depending on the coverage of the relationship;
    - Otherwise, the mapping is structural or non-existent and *MatchClass*∈ **Table** 2.4∪{*Equal, Disjoint*} ;

- *Exp* stands for the mapping expression that translates and further specifies the previous tuple components. It can be written using a finite set of binary operators derived from the mathematical symbols associated with the mapping types and classes (e.g. "=, ~, ⊆, ⊇, ⊥, +,-, ×, ÷, *concatenate, split*").

This mapping tuple which represents *θ(a,b),* can also be used to generate a transformation function τ, where $\tau(a, \theta) = b$, being$(a, b) \in MElems$. This will enable an automatic data transformation and exchange between two organizations working with different information models, thus achieving an interoperable state among them or supporting the recovery from any harmonization breaking situation.

In order to enable the envisaged traceability to support intelligence and sustainability, it is required not only to have the morphism formalised using a textual tuple, but also to store it in a parseable and structured KB. With it, every mapping between models or ontologies of business partners can be stored and accessed by their local systems. This allows communities to build systems with reasoning capabilities able to understand each other's representation format, without having to change their data and schema import or export processes [29].

2.3.3 **KU Conclusions**

Enterprises are demanded to collaborate and establish partnerships to reach global business and markets. However, due to the different sources of models and semantics, organizations are experiencing difficulties exchanging vital information electronically and seamlessly, even when they operate in related business environments. This situation is even worst in the advent of the evolution of the enterprise systems and applications, whose dynamics result in increasing the interoperability problem due to the continuous need for model adjustments and semantics harmonization. To contribute for a long term stable interoperable enterprise operating environment, the author proposes the integration of traceability functionalities in the information systems as a way to support such desired stage. Either data, semantic, and structural mappings between partner enterprises in the complex network should be modelled as tuples and stored in a communication KB with reasoning capabilities, thus allowing tracing, monitoring and supporting the stability maintenance of a system's interoperable state.

2.4 Knowledge Maintenance (KMa)

Systems data and knowledge change over time, and even rules, which were valid once eventually, become riddled with exceptions, and a specialist who does not adapt will find his work to become without value [140]. Thus, information systems must be able to continuously deal with such continuing changes explicitly in a formal dynamic Knowledge Maintenance (KMa) establishment. KMa is focused on the KB improvement to actively be updated, monitored accordingly to the knowledge evolution of its related domain.

In additional, KMa has become a challenging activity for most competitive business organisations. There is growing recognition in the business community about the importance of knowledge as a critical resource for organisations [141][142][143]. Individuals and companies are obliged to focus on maintaining and enhancing their knowledge asset in order to innovate [144][143] and survive in the current competitive markets.

The concepts meanings are the main focus of the knowledge acquisition and maintenance procedures. Thus, semantic relatedness techniques have a main role in the KMa establishment. Furthermore, in line with this, Witherell stated that, "leveraging works in semantic relatedness methods can be developed to give 'intelligent' knowledge frameworks the ability to automatically adapt within fluid environments" [145]. In complement to this, Witherell concluded that those knowledge frameworks are founded on the ability of ontologies to explicitly represent and structure information and the ability of semantic relatedness techniques to interpret relationships formed between concepts in such kind of knowledge structure [145]. This also complements the idea that semantic adaptation ability increases the intelligence level of an information system dealing with interoperability.

Since this dissertation addresses artificial (computational) systems, this section starts by introducing some statements about AI and its parallelism to neuroscience addressing how human brain deals with its semantic memories and learning. Then, human based learning techniques; OL and operation research methods are introduced that together with some machine learning techniques could be used to facilitate knowledge systems maintenance.

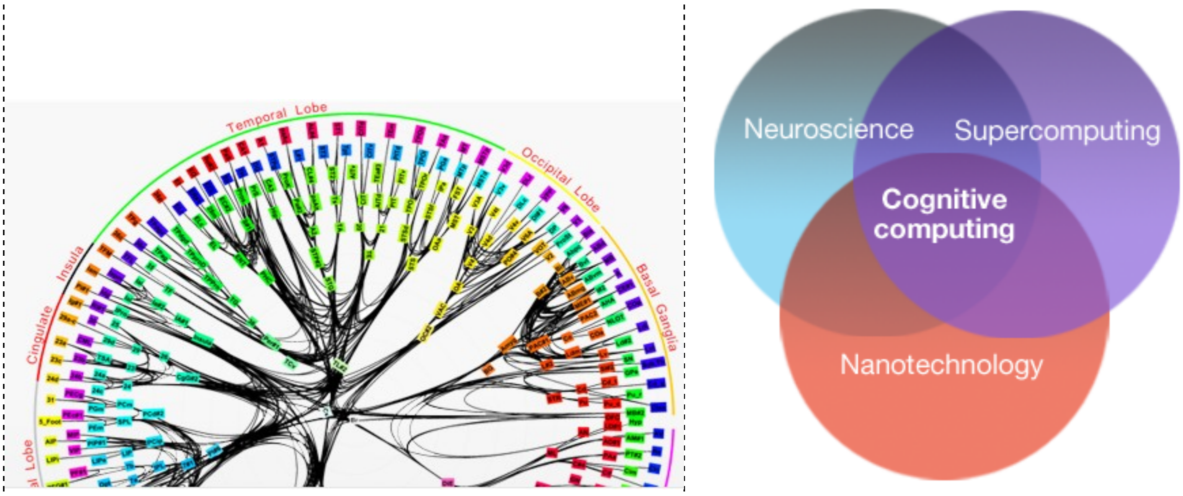
2.4.1 **Artificial Intelligence and Neuroscience**

Since its conception in the mid-1950s, AI with its great ambition to understand and emulate intelligence in natural and artificial environments alike is now a truly multidisciplinary field that reaches out and is inspired by a great diversity of other fields.

AI has dreamt about for a long time to simulate in computer an entire biological brain or the creation of new life forms from manipulations of cellular and genetic information. The scope for AI in neuroscience and systems biology is extremely wide. Berrar et al. in [146] investigated the standing of AI in relation to neuroscience and systems biology and provided an outlook at new and exciting challenges for AI in these fields.

Neuroscience is a tremendously complex field with many subfields (e.g., [147]) and offers a wide range of opportunities for AI. Neural signalling, for instance, investigates how neural systems acquire, coordinate, and disseminate information. Knowledge about these processes is fundamental to understanding brain pathologies, but also for the development of novel approaches to diagnosing and treating such problems. AI systems can benefit from such, an understanding in the field of ANN. Other application areas that AI can benefit are pervasive/ubiquitous computing [148] and autonomic computing [149]. This pursues the idea of having networked communication systems able of being autonomously controlled, in a sense that it can manage themselves as a kind of administrators.

Along this line, brain-like computers are another exciting area in neuroscience. For instance, IBM has engaged into an activity researching brain-like computers and the newly coined term cognitive computing encapsulates the key idea to engineer mind-like intelligent machines by reverse engineering the structure, dynamics, function and behaviour of the brain [150]. This has been addressed through the program Systems of Neuromorphic Adaptive Plastic Scalable Electronics (SyNAPSE) (Figure 2.12 a)). Such program intends to combine principles from nanoscience, neuroscience and supercomputing as part of a multi-year cognitive computing initiative (Figure 2.12 b)). The initial phase of SyNAPSE developed nanometre-scale electronic synaptic components capable of adapting connection strength between two neurons in a manner analogous to what is performed by the human brain, which autonomously process information in complex environments by automatically learning relevant and probabilistically stable features and associations [151].



a) b)

Figure 2.12 – a) Artificial brain-like connections; and b) Cognitive computing principles [152]

Current programmable machines work superbly with numbers but are limited describing and processing information from particular environments, like in human faces or objects recognition. Thus, this initiative is developing cognitive computers, whose use advanced algorithms and silicon circuitry to be able to learn through experiences, find correlations, create hypotheses, remember and learn from the outcomes [152].

Semantic Memory on the brain

Semantic memory, also called conceptual knowledge, is the aspect of human memory that corresponds to general knowledge of objects, word meanings, facts and people, without connection to any particular time or place [153]. This provides to humans the ability to, given a concept, be able to create its mental visualization and also to see the most important associations that the individual has to the referred concept. Although this kind of memory may depend on individual's experience, which is mostly shared in a given culture.

Patterson et al. identified two theoretical positions regarding the neuroanatomical distribution of the cortical semantic network and schematic models based on these views (Figure 2.13) [154]. The distributed-only theoretical position view proposes that the semantic network in the brain is composed by those regions, along with all the diverse connections between them, represented by the green lines. The flow of activation through this network can be 'gated' by a representation of the current task (right-hand panel): for instance, if the task is to name a line drawing of a familiar object, activation will flow from a representation of object's shape to a representation of its name. Associations between different pairs of attributes are encoded along different neuroanatomical pathways.

However, there are several cases of individuals whom suffered brain damages that would affect that regions but could recover. This may lead to that probably the entire neural basis of semantic memory might have a different approach. Figure 2.13 (part b – Distributed-plus-hub view) suggests that in addition to distributed-only view proposed representation, there are connections (shown as red lines) between the various shapes representations and a modal hub. It's in this hub where the associations between different pairs of attributes (such as shape and name or shape and action, etc.) are processed by a common set of neurons and synapses regardless of task representation. Figures right-hand side shows the corresponding convergent architecture of these views [154]. Thus, there is an extra path in semantic representation in this theoretical approach, which could explain the recovering status of the mentioned brain damage example.

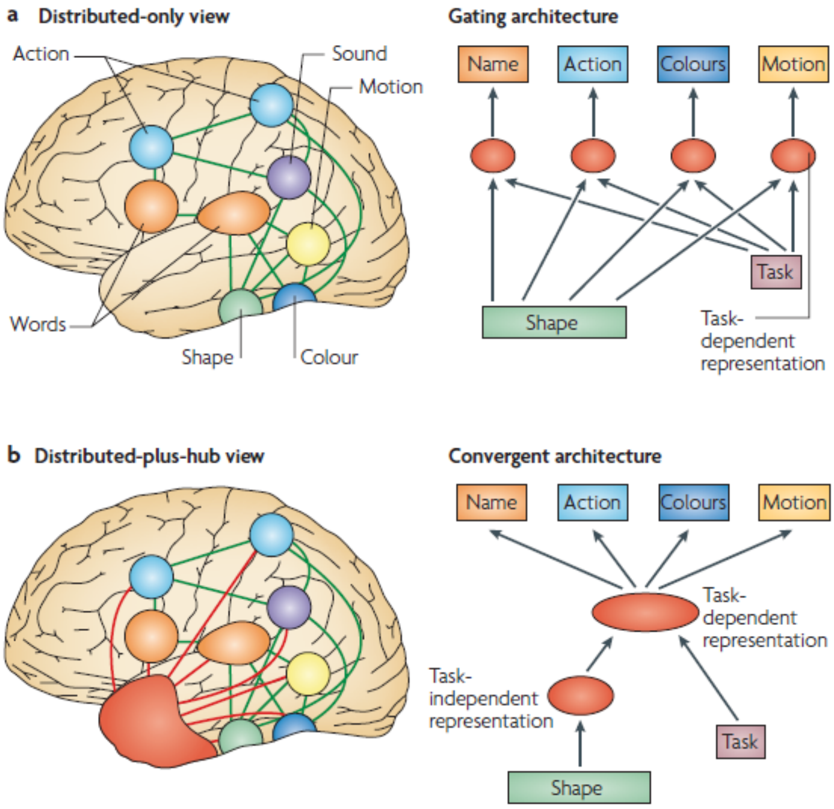


Figure 2.13 – Two theoretical positions regarding the neuroanatomical distribution of the cortical semantic network and schematic models based on these views [154].

Brain Learning

Neuroscience has reached the stage where it is possible to understand how parts of the brain actually work, by combining approaches from many disciplines. Evidence on the connections and internal connectivity of each brain region, of the biophysical properties of single neurons [155], on what is represented by neuronal activity in each brain region, and on the effects of lesions, all provide the foundation for a computational understanding of brain function in terms of the neuronal network operations being performed in each region [156]. Crucial brain systems to understand are those involved in memory, but in addition, learning mechanisms are at the heart of how the brain processes information, for it is by modifying the synaptic connection strengths (or weights) between neurons that useful neuronal information processors for most brain functions, including perception, emotion, motivation, and motor function, are built [157].

In addition to prove that the mentioned learning mechanisms are directly related to the synaptic relation strengths modification it is presented in the following, some facts about a study made by Patterson et al. in [154] about learning using animals. Patterson et al. used basic emotions as a facilitator for learning. Thus, emotions were defined in this case, due to the animal's use, as states elicited by rewards and punishers [158][159][160]. A reward is anything for which an animal will work. A punisher is anything an animal will work to escape or avoid. Rewards and punishers can be more formally defined as instrumental reinforcers, i.e. stimuli or events which, if their occurrence, termination, or omission is made contingent upon the making of a response, alter the probability of the future emission of that response. Some stimuli are unlearned, or primary, reinforcers (e.g. pain or the taste of food if the animal is hungry), whereas others may become reinforcing by learning, through their association with primary reinforcers, thereby becoming secondary reinforcers. This type of learning may thus be called stimulus-reinforcement association learning [157].

Accordingly to this, Hebbian theory about learning states that simultaneous activation of cells leads to increases in synaptic strength between those cells [161]. These weights are adjusted in order to better represent the relationship between two cells. Transporting it into ANN, it can be a method of determining how to increase/decrease weights between two neurons.

2.4.2 **Artificial Learning Related Techniques inspired on Human Learning**

In the pursuit of artificial intelligent semantic relatedness techniques/methods able to facilitate semantic adaptation ability, author introduces some of artificial learning related techniques inspired on human learning. Consequently, ANN, FL and OL are presented.

Artificial Neural Networks

An ANN is an information-processing paradigm that is inspired by the way biological nervous system, such as brain, process information [162]. It can be represented by a directed graph consisting of nodes with interconnecting synaptic and activation links [163], which is composed through neurons that are represented by a set of linear synaptic links, an externally applied bias, and a possibly nonlinear activation link. The basic computational element of an ANN is a neuron also often called node or unit. The synaptic links of a neuron weight their respective input signals. The weighted sum of the input signals defines the inducted local field of the neuron in question, and the activation link squashes the induced local field of the neuron to produce an output. Such weights can be adjusted in a learning process to put the ANN answering in a pre-determined form. The neuron computes some function f of the weighted sum of its inputs: $y_i = f\ (\ \sum w_{ij}\ y_j\ )$. Its output, in turn, can serve as input to other units. Where the weighted sum $\sum_j w_{ij}\ y_j$ is called the net input to unit *i*, often written net_i. and w_{ij} refers to the weight from unit *j* to unit *i*. The function *f* is the unit's activation function that in a simple case, *f* can be a linear unit, which its output is just its net input (Figure 2.14 a)) [164].

An ANN is like human beings, it learns by example, therefore it can be configured for a specific application, such as data classification or patterns recognition to detect trends that are too complex to be noticed by humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information to what it was given to analyse. All this process is similar to the learning in biological systems, which involves adjustments to synaptic connections that exist between the neurones.

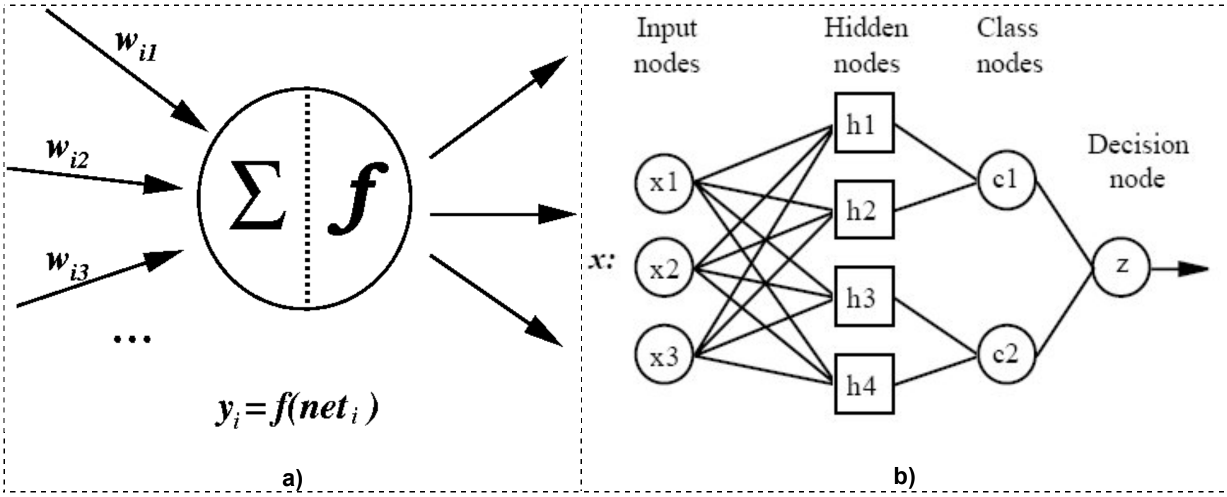


Figure 2.14 – a) A Neuron function [164]; b) A Probabilistic Neural Network [165]

There are various types of ANNs developed for specific objectives. In 1990, Donald F. Specht proposed a method to formulate the weighted-neighbour method in the form of a neural network, which resulted in the "*Probabilistic Neural Network*" that is illustrated in Figure 2.14 b). However, there are many other types like: Multilayer Perceptron Networks, General Regression Neural Networks, Radial Basis Function Networks, Cascade Correlation, Functional Link Networks, Kohonen networks, Gram-Charlier networks, Learning Vector Quantization, Hebb networks, Adaline networks, Heteroassociative networks, Recurrent Networks and Hybrid Networks [165].

Fuzzy Logic

The concept of FL was conceived by Lotfi Zadeh, a professor at the University of California at Berkley. He reasoned that people do not require precise, numerical information input, and yet they are capable of highly adaptive control. If feedback controllers could be programmed to accept noisy, imprecise input, they would be much more effective and perhaps easier to implement [166]. There are several unique features that make FL a particularly good choice for many control problems. The output control is a smooth control function despite a wide range of input variations. FL can control nonlinear systems that would be difficult or impossible to model mathematically. This opens doors for control systems that would normally be deemed unfeasible for automation.

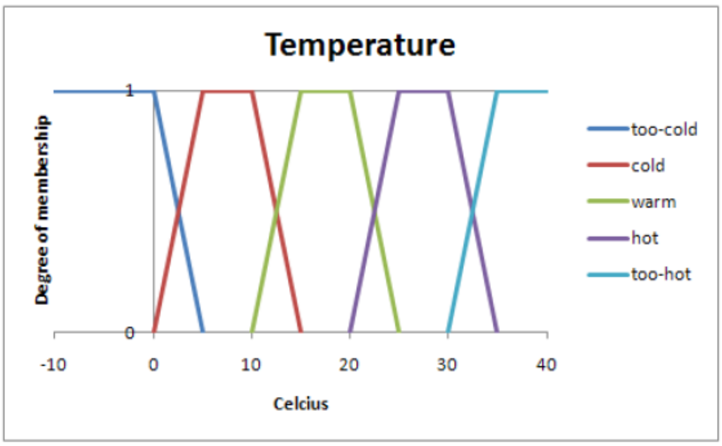


Figure 2.15 – Fuzzy Membership Functions [167]

FL incorporates a simple, rule-based "IF X AND Y THEN Z" approach to a solving control problem rather than attempting to model a system mathematically. A membership function is then used to quantify a linguistic term like temperature [167]. For example, according to Figure 2.15, a temperature value can be considered as "cold" and "too-cold" at the same time, with different degree of memberships. In this case, these membership functions use a trapezoidal form, but there are other forms as triangular, trapezoidal, piecewise linear, gaussian, or singleton. Taking as reference the matrix of target values presented at *Table 2.5*, a rule example for this case could be "IF (temperature is "cold" OR "too-cold") AND (target is "warm") THEN (add "heat" to the process)".

temperature/target	too-cold	cold	warm	hot	too-hot
too-cold	no-change	heat	heat	heat	heat
cold	cool	no-change	heat	heat	heat
warm	cool	cool	no-change	heat	heat
hot	cool	cool	cool	no-change	heat
too-hot	cool	cool	cool	cool	no-change

Table 2.5 – A Fuzzy Matrix Example [167]

Despite being largely data-driven, machine learning is not a knowledge-free methodology. Instead, successful learning requires a priori background knowledge and a proper modelling of the data (which is often overlooked) and the underlying hypothesis space (type of model), and FL can be very useful in this regard (incorporating expert knowledge into the learning process)! FL is then a kind of mathematical framework complementary to statistics and probability for formalizing (human-like) approximate reasoning [168].

Ontology Learning

The main idea is to find a system able to learn, but also able to represent and manage complex inputs as concepts like humans do. Despite ANN being able to learn and FL to respond like humans reasoning to complex systems/situations, it is needed to analyse other learning techniques that are also able to manage concepts. To structure and manage data (concepts) in a way that machines can handle ontologies can be used, even more when ontologies structure knowledge in a taxonomic way also similar to humans. Thus, its integration with some machine learning abilities, leads to the appearance of the OL concept. OL is mainly associated by the research community to the process of discovering ontological knowledge from various forms of data, as text, using an automatic or semi-automatic process [169]. However, a full automatic machine learning knowledge acquisition was not yet reached, but its actual solutions are considered a powerful tool to assist in the management of the ontologies [51].

OL provides the ability to not only discover ontological knowledge at a larger scale and a faster pace, but also mitigate human-introduced biases and inconsistencies. In additional, it can be used also as a support to the refinement and expansion of existing (reference) ontologies (that could have been built following a traditional basis with human supervision) by incorporating new knowledge [169] in an automatic way. Thus, it has to be said that all the represented knowledge was built in a semi-automatic process. However, this approach represents a more confident KMa process, because it has a stable reference background knowledge, which will receive new knowledge in automatic way without changing the basis of its overall represented knowledge.

To facilitate OL related techniques as ontology construction and discover ontological knowledge, machine learning is commonly used [169][51]. However, the majority of these OL techniques are unsupervised resulting in a higher employment of statistical techniques [169].

Operation research methods

Operations research is an interdisciplinary mathematical science, as it employs techniques from other mathematical sciences, such as mathematical modelling, statistical analysis, and mathematical optimization that focus on the effective use of technology by organizations, providing solutions to complex decision-making problems [170], and in the pursuit of its improvement and efficiency [171].

A problem in the real world is modelled, usually in mathematical terms, then mathematical techniques, together with data analysis and computational algorithms are applied, in order to find ways to do the job better [172]. Operation research methods are used to support such problems resolution. Thus, they were introduced in OL to provide a statistical analysis of ontologies concepts and relations to facilitate complex decision-making effectiveness. Markov Chains (MC) and Bayesian Networks (BNs) are two probability models in the field of operations research, which researchers have applied to support OL.

Markov Chains

Markov logic is a novel language [173] that provides the capability of joining in the same representation statistical and relational aspects [174], which result is a probabilistic graphic model called Markov networks or MC. However, when there is a statistical relational learning technique that combines first order logic and MC the result is the Markov Logic Networks (MLN). The combination of these two techniques (statistical relational learning and MC) gives us the flexibility of knowledge representation as well as the ability to handle uncertainty [175].

A MLN can be built by attaching weights to first-order logic formulas as in a case, where the higher the weight, the bigger is the difference between a world that satisfies the formula and one that does not. A network like this can be used as a template to construct a kind of Markov networks, providing the full expressiveness of probabilistic models and first-order logic [173].

Given a set of constants (i.e., individuals) of the domain and an interpretation, the groundings of the formulas in an MLN can generate a Markov network by adding a variable for each ground atom, an edge if two ground atoms appear in the same formula, and a feature for each grounded formula. The probability distribution of the network is defined as $P(X = x) = \frac{1}{Z} exp(\sum_{i=1}^{F} wi\, ni\ (x))$, where F is the number of formulas in the MLN, $n_i(x)$ is the (binary) number of true groundings of F_i in the world x, W_i is the weight of F_i, and Z is a normalizing constant. Formulas' weights can be learned generatively from example data by maximizing the pseudo-log-likelihood [174] of that data, while efficient inference can be done using approximate inference algorithms, such as the MC-SAT[16] [174].

A MC like the one presented in Figure 2.16, consists of a finite number of states (which may be

[16] MC-SAT is an Markov chain Monte Carlo inference algorithm designed to deal efficiently with probabilistic and deterministic dependencies – Retrieved from the web at October 2011 : http://alchemy.cs.washington.edu/api/html/classMCSAT.html#_details

represented by circles) and paths between states with assigned probabilities (which may be represented by arrows). The probabilities on all arrows leaving a state should add to one. Consider a walk on the set of states of a MC where each step follows an arrow in it. Choosing which state to go to next depends only on your current state and the probabilities on the arrows leaving your current state. In other words, each step is independent of all steps that happened before you landed on your current state (the “history" of that state) [176].

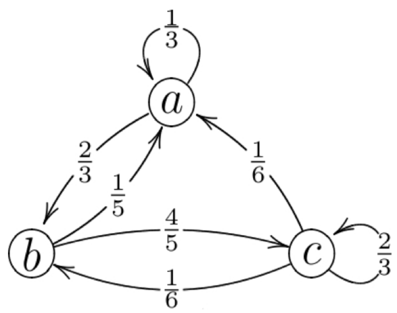


Figure 2.16 – A Markov Chain [176]

There are several areas where Markov process has an important rule, such as in statistic, marketing, genetic, computer vision, diagnostic and troubleshooting, software debugging, speech recognition and understanding algorithms, Internet, musical composition, etc. [177]. Even not knowing, everyday millions of Internet users use Markov in some navigation patterns and especially in Google “RankPage”. The PageRank of a webpage is the probability to be at page *i* in the stationary distribution on the following Markov chain on all (known) webpages [178]. If N is the number of known webpages, and a page *i* has k_i links then it has transition probability $\frac{\alpha}{ki} + \frac{1-\alpha}{N}$ for all pages that are linked to and $\frac{1-\alpha}{N}$ for all pages that are not linked to. The parameter α is taken to be about 0.85. Markov models have also been used to analyse web navigation behaviour of users. A user's web link transition on a particular website can be modelled using first or second order Markov models and can be used to make predictions regarding future navigation and to personalize the web page for an individual user.

Another interesting application of Markov networks is about to build a probabilistic scheme for ontology matching [179], one of this cases is iMatch. iMatch is a novel probabilistic scheme for ontology matching, where a Markov network is constructed on the fly according to the two input ontologies; evidence from first-line matchers is introduced, and probabilistic reasoning is used to produce matchings [179]. Likelihood knowledge can be introduced to improve the associated MLN learning process.

MLN is also used by OL techniques. Drumond in [180] presents an OL framework supported by MLN. MLN were used specifically in PRECE (Probabilistic RElational Concept Extraction) and PREHE (Probabilistic RElational Hierarchy Extraction) techniques. These techniques are respectively related to the concepts extraction from textual sources and then for the extraction of taxonomic relations between those concepts [180]. In this same work example, some initial knowledge-based on

Wordnet[17] is introduced related to the text pre-processing to act as initial knowledge to put the used MLN working for text corpus identification.

In a kind of short conclusion it is needed to state that MLN needs the support of learning algorithms, whose can be of various types as conjugate gradient algorithm; pseudo-likelihood; inductive logic programming; etc. to readapt its related statistics. In additional, they need some initial knowledge (e.g. semantic relations with some statistical information) to start its learning process.

Bayesian Networks

A BN is a Directed Acyclic Graph (DAG) with attached local probability distributions [181]. Nodes in the graph represent random variables (corresponding to attributes, features etc.). Each random variable has a mutually exclusive and exhaustive set of values (states). Edges in the graph represent direct interdependences between two random variables. BNs, uses then the following operators: for any pair of nodes x and y, if x and y are adjacent, the edge connecting them can be either deleted or reversed. If x and y are not adjacent, an edge can be added in either direction. All operators are subject to the constraint that a cycle cannot be formed [182]. BNs consist of two sort of knowledge:

- qualitative knowledge that describes interdependencies by means of directed graph;
- quantitative knowledge that captures relations among random variables by means of Conditional Probability Tables (CPT). An advantage of BNs compared to other uncertainty representation formalisms is the possibility to model complicated mutually related phenomena in quite a tractable way.

Figure 2.17 represents a BN graph, which nodes have CPT representing random variables, for having a wet grass situation depending on rain or sprinkler and its relation to a cloudy day. Thus its arrows connect such nodes to represent the influence between them [183].

BNs provide a means of capturing existing knowledge about a domain, learning the stochastic properties of that domain and thereby adjusting its model. This adjust can be for instance related to the ontologies mapping establishment. BNs are currently being exploited also for estimating effects of different types of behaviour and as support for human or automated decision tasks. Some sample applications include using BNs to reduce power consumption of machines with reference to user behaviour [184], to diagnose faults in industrial processes [185] or to monitoring and manipulating cause and effects for modelled systems as disparate as the weather, disease and mobile telecommunications networks [186].

[17] WordNet® is a large lexical database of English. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations. Retrieved from the web at October 2011: http://wordnet.princeton.edu/

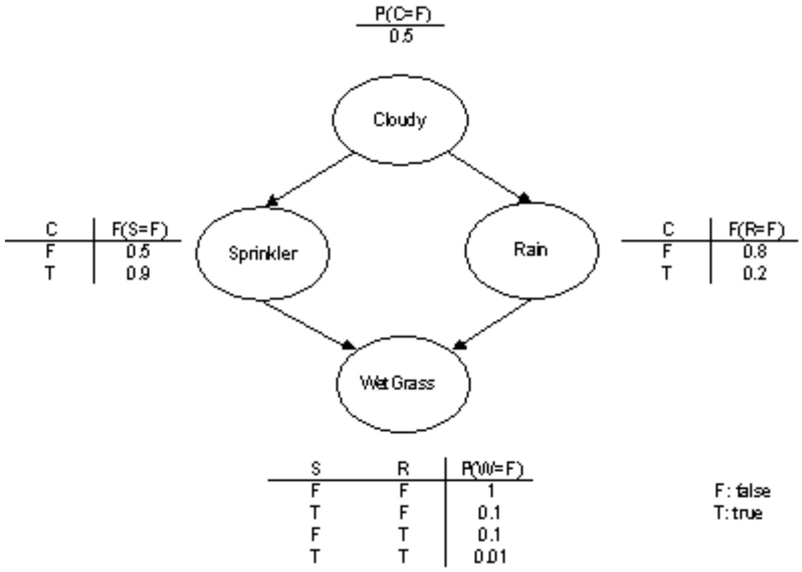


Figure 2.17 – A Wet Grass BN [183]

Colace et al. in [187] presented a method for learning curricula ontologies using an approach based on BNs. This approach collects data drawn from real learning experiences and by the automatic analysis of students' learning performances to check the courses' ontology and propose corrective actions. It also provides an ontological basis for determining learning paths to personalize learning, which is done through the matching between the ontology and the BN. The proposed tool allows effective tutoring and adaptations of the learning process to the demands of students.

This research, which merges ontologies with BN conducted to the appearance of the **BayesOWL** framework. It augments and supplements OWL for representing and reasoning with uncertainty based on BNs [188][189]. This framework consists of three key components: 1) a representation of probabilistic constraints as OWL statements; 2) a set of structural translation rules and procedures that converts an OWL taxonomy ontology into a BN DAG; and 3) a Simplified D-IPFP [190] based on 'Iterative Proportional Fitting Procedure' (IPFP)[18] that incorporates available probability constraints into the CPT of the translated BN. The translated BN, which preserves the semantics of the original ontology and is consistent with all the given probability constraints, can support ontology reasoning, both within and cross ontologies, as Bayesian inferences, with more accurate and more plausible results [191]. This framework was transformed in a Java-based tool to extract taxonomy from OWL ontology, translate it into Bayesian Nets and integrate uncertainty knowledge into result BNs. However its developing community is not so much active.

Due to the fact that uncertainty is ubiquitous, any representation scheme intended to model real-world actions and processes must be able to cope with the effects of uncertain phenomena [192]. Costa et

[18] Iterative Proportional Fitting Procedure (IPFP) is a mathematical procedure that iteratively modifies a probability distribution to satisfy a set of probability constraints while maintaining minimum Kullback-Leibler distance to the original distribution [190]. Where Kullback-Leibler distance is a natural distance function from a "true" probability distribution, p, to a "target" probability distribution, q. It can be interpreted as the expected extra message-length per datum due to using a code based on the wrong (target) distribution compared to using a code based on the true distribution. – Last sentence retrieved from the web at November 2011: http://www.csse.monash.edu.au/~lloyd/tildeMML/KL/

al. created the **PR-OWL** open research work, which aim is to establish a framework to enable the full support for uncertainty in the field of ontology engineering and, as a consequence, for the Semantic Web [193]. Its approach involves augmenting OWL semantics to allow probabilistic information to be represented via additional mark-ups. Then, a set of structural translation rules is defined to convert this probabilistically annotated OWL ontology taxonomy into a DAG of a BN. Finally, the BN is completed by constructing CPT for each node in the DAG.

The development of such research work leaded to the conception of the probabilistic ontology concept. Costa et al. defined it in [193]: "*A probabilistic ontology is an explicit, formal KR that expresses knowledge about a domain of application. This includes: (i) Types of entities that exist in the domain; (ii) Properties of those entities; (iii) Relationships among entities; (iv) Processes and events that happen with those entities; (v) Statistical regularities that characterize the domain; (vi) Inconclusive, ambiguous, incomplete, unreliable, and dissonant knowledge related to entities of the domain; and (vii) Uncertainty about all the above forms of knowledge*".

As mentioned by Drudzel et al in [194], the task of building the structure and assigning the probability distributions of a BN is complex and knowledge-intensive. Moreover, it could be said the same for MLN and consequently to the probabilistic ontologies about such probability assignments. All of these probability models require the identification of relevant statistical variables in the application domain, the assignment of initial probability distributions and in the case of using BN, the specification of dependency relations between these variables [194]. Even though, these probability models still need support of learning algorithms to conduct the learning process as in readjusting their associated statistics. Consequently a machine learning literature review is presenting in the following since it could help on defining these supporting algorithms.

2.4.3 Machine Learning

Machine learning is a scientific discipline concerned with the design and development of algorithms that allow computers to evolve behaviours based on empirical data, such as from sensor data or databases [195]. It studies computer learning algorithms to do stuff. This learning must be automatically done without human intervention or assistance. Machine learning paradigm can be viewed as "programming by example", learning to do better in the future, based on what was experienced in the past [196].

Recent advances in machine learning make possible to design efficient prediction algorithms for data sets with huge number of parameters. For example one loan company has used a statistical decision procedure based on machine learning since it discovered that it produces around two-thirds of correct predictions. It has been implied since then, not only because of these rules success rate of the loan decisions, but also because the company found them attractive to be used to explain to applicants the reasons behind the decision [197].

Machine learning market applications are endless. They can be applied to optical character recognition, face detections, spam filtering, topic spotting, spoken language understanding, medical diagnosis, fraud detection, weather prediction, etc. Moreover, there are some areas where machine learning techniques reached a level of performance equal or even greater than human experts. For example in astronomy, machine learning has been used to develop a fully automatic cataloguing system for celestial objects that are too faint to be seen by visual inspection [197].

Research community has used a few various sorts of machine learning categorisations like: model types; reasoning types; general types of machine learning; learning techniques/task types etc. The author describes in the following two of these categorisation types: the general machine learning and learning techniques/tasks types.

Most of the researchers use three general machine learning classification types: the supervised; the unsupervised and the reinforcement learning. But, there are others that use more than three or even a mix of the above ones to better classify some approaches (e.g. semi-supervised learning, which combines both supervised and unsupervised characteristics) [198].

The **supervised learning** aggregates algorithms that generate a function that maps inputs to desired outputs. It consists of examples which include both the inputs and the desired outputs, thus enabling it to learn the function. The learner should then be able to generalize from the presented data to unseen examples.

The **unsupervised learning** represents algorithms that are able to fit in a model a set of observed characteristics of a predetermined set of examples/objects. Then based on that model, the algorithm is able to classify other inputs as similar/equal or not to the modelled ones. For example, a clustering algorithm would be a form of unsupervised learning.

Reinforcement learning learns how to act given an observation of the world. Every action has some impact in the environment, and the environment provides feedback in the form of rewards that guides the learning algorithm [195].

Witten et al. states in [197] that the learning techniques/task types are basically four in use: Classification; Association; Clustering; and Numeric Prediction.

Classification learning

Classification learning is a learning scheme for categorizing unseen examples into predefined classes based on a set of training examples. The learning algorithm generates a set of classification rules from a complete set of independent examples of instances and their corresponding categories, and then the generated rules are used to predict the classes or categories of novel instances. The purpose of classification learning is to predict the classes of instances [199]. Classification learning is sometimes also called supervised, because, in a sense, the method operates under supervision by being provided with the actual outcome for each of the training examples [197]. This outcome is called the class of the example [200].

In classification learning if the target variable is categorical, then a classification tree is generated. To predict the value (category) of the target variable using a classification tree it is used the values of the predictor variables to move through the tree until reach a terminal (leaf) node, and then it is predicted the category shown for that node. An example of a classification tree is shown in Figure 2.18. The target variable is “Species”, the species of Iris. From the tree it could be seen that if the value of the predictor variable “Petal length” is less than or equal to 2.45, the species is “Setosa”. If the petal length is greater than 2.45, then additional splits are required to classify the species [165].

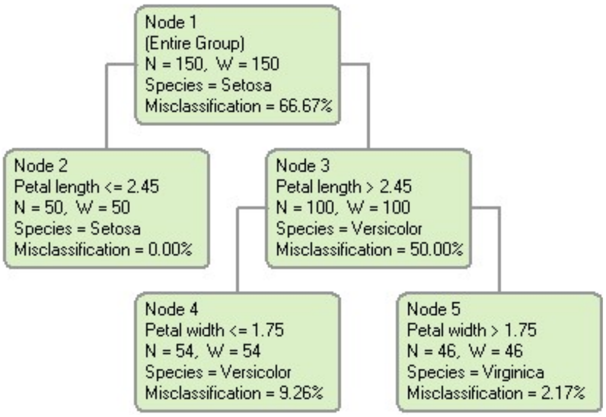
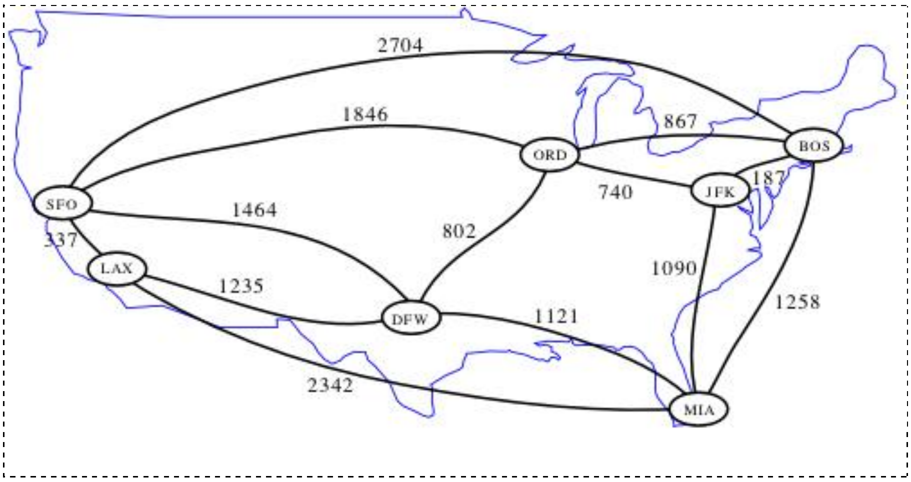


Figure 2.18 – A Classification tree [165]

One of the classification learning algorithms is the 1R procedure. Called 1R for 1-rule, it generates a one-level decision tree expressed in the form of a set of rules that all test one particular attribute. 1R is a simple, cheap method that often comes up with quite good rules for characterizing the structure in data. It turns out that simple rules frequently achieve surprisingly high accuracy. This simple technique often gives a useful first impression of a dataset. It is extremely fast and can be applied to very large quantities of data [197].

Instance-based learning is a kind of classification learning. In instance-based learning training examples are stored verbatim, and a distance function is used to determine which member of the training set is closest to an unknown test instance [197]. Once the nearest training instance has been located, its class is predicted for the test instance. Some varieties of instance-based learning deal only with ratio scales because they calculate the “distance” between two instances based on the values of their attributes [197]. If the actual scale is ordinal, a numeric distance function must be defined [197].

Although there are other possible choices, most instance-based learners use Euclidean distance. The distance between an instance with attribute values ${a_1}^{(1)}, {a_2}^{(1)}, \dots, {a_k}^{(1)}$ (where k is the number of attributes) and one with values ${a_1}^{(2)}$, ${a_2}^{(2)}$, ... , ${a_k}^{(2)}$ is defined as $\sqrt{({a_1}^{(1)} - {a_1}^{(2)})^2({a_2}^{(1)} - {a_2}^{(2)})^2 + \cdots + ({a_k}^{(1)} - {a_k}^{(2)})^2}$. When comparing distances it is not necessary to perform the square root operation; the sums of squares can be compared directly [197]. However, the shortest distance calculation can be supported by the shortest path algorithms application as author describes in a prototype development at chapter 6. There are several algorithms for shortest path resolution like Alpha-beta pruning; A*; B*; Bellman–Ford algorithm; Dijkstra's algorithm; Floyd–Warshall; etc. Due to its availability in Java, author used Dijkstra’s algorithm for supporting semantic distance calculation. Dijkstra’s algorithm, published in 1959 [201], is a graph search algorithm to produce the shortest path in a graph with nonnegative edge path costs. Figure 2.19 illustrates a weighted graph whose vertices represent major US airports and whose edge weights represent distances in miles. This graph has a path from JFK to LAX of total weight 2777 (going through ORD and DFW). This is the minimum weight path in the graph from JFK to LAX [202]. This information can be used, for instance, to classify ORD and DFW as “on the way” airports between JFK and LAX. This information represents a learning result from the shortest path algorithm application in the presented weighted graph.



Legend:

BOS – Airport at Boston
DFW – Airport at Dallas
JFK – Airport at New York
LAX – Airport at Los Angeles
MIA – Airport at Miami
ORD – Airport at Chicago
SFO - Airport at San Francisco

Figure 2.19 – Weighted Graph representing distances in miles between US airports [202]

Association learning

Data can be mined to identify associations. Introduced in 1993 [203] the task association rule mining has received a great deal of attention till nowadays where the meaning of such rules is still one of the most popular pattern discovery methods in knowledge discovery data [204]. Association rules differ from classification rules as they can predict not just the class but also any attribute and more than one attribute’s value at a time [197]. Association rule learning is a popular and well researched method for discovering interesting relations between variables in large databases [205]. An example of an association rule about grocery purchases is “*People who buy diapers tend to buy beer* “[206]. This is the reason for the higher existing number of association rules than classification rules. Association rules usually involve only nonnumeric attributes [197]. OL also focuses on association learning [169]. The generalized association-rule-Learning algorithm extends its baseline by aiming at descriptions at the appropriate taxonomy level. For example, “snacks are purchased together with drinks” [51]. Figure 2.20 presents a supermarket basket example statistical analysis to reach a set of association rules. From this example it was defined that a basket with a product A, will have also a product C with a confidence rate of 2/3.

Apriori is an iterative association rule learning (seminal) algorithm for finding frequent item sets using candidate generation [208]. It is characterized by finding frequent item sets from a transaction dataset and deriving association rules. Finding frequent item sets (item sets with frequency larger than or equal to a user specified minimum support), is not trivial because of its combinatorial explosion. Once frequent item sets are obtained, it is straightforward to generate association rules with confidence larger than or equal to a user specified minimum confidence [209]. The introduction of this technique boosted data mining research and its impact is tremendous. The algorithm is quite simple and easy to implement [210].

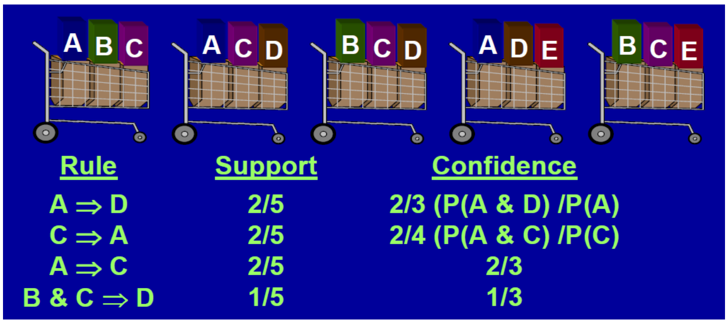


Figure 2.20 – Supermarket association rule example [203][207]

Clustering

Clustering technique is an unsupervised learning task to learn a classification from the data [200], usually applied to group items that seem to fall naturally together [197] instead of requiring a predefined classification [200]. In clustering the classes are not predefined but are unknown at the point of learning, and the learning task is to define and identify the classes in the database [199]. The challenge is to find these classes/clusters and assign the instances to them [197]. Data items are grouped according to logical relationships, for example in consumer preferences, data can be mined to identify market segments or consumer affinities. It is a common technique for statistical data analysis [195].

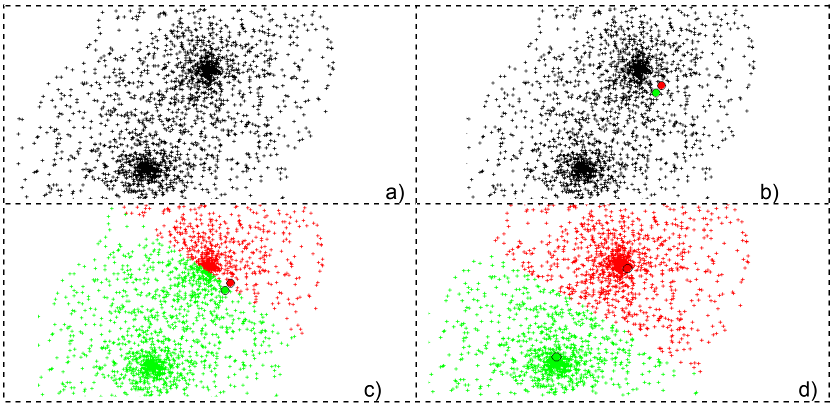


Figure 2.21 – Clustering learning using K-means algorithm

The classic clustering technique is called k-means. All instances are assigned to their closest cluster centre according to the ordinary Euclidean distance metric. The k-means clustering algorithm usually requires several iterations, each involving finding the distance of k cluster centres from every instance to determine its cluster [197]. It remains the most widely used partitioned clustering algorithm in practice. The algorithm is simple, easily understandable and reasonably scalable, and can be easily modified to deal with streaming data. Note that each iteration needs $N \times k$ comparisons, which determines the time complexity of one iteration. The number of iterations required for convergence varies and may depend on N, but as a first cut, this algorithm can be considered linear in the dataset size. One issue to resolve is how to quantify "closest" in the assignment step. The default measure of closeness is the Euclidean distance, in which case one can readily show that the non-negative cost function will decrease whenever there is a change in the assignment or the relocation steps, and hence convergence is guaranteed in a finite number of iterations [210]. Figure 2.21 illustrates some steps of an example using k-means generated at [211]: a) presents the initial elements/objects to cluster; b) randomly it is generated two initial group centroids (red and green points); c) presents the assignment of each object to each group that has closest to each centroid (first iteration); then it is recalculated iteratively the positions of the centroids until they stop changing, which in this specific example happened when they reach the position shown at d).

Numeric Prediction

Numeric prediction or regression, classes are composed not of discrete categories but of continuous numeric values, but otherwise regression learning uses techniques very similar to classification learning and is sometimes considered a subtype of classification learning [199]. Numeric Prediction is a variant of classification learning in which the outcome is a numeric value rather than a category. Linear regression is a natural technique to be considered. Its idea is to express the class (x) as a linear combination of the attributes $(a_1, a_2, \ldots, a_k)$, with predetermined weights$(w_0, w_1, \ldots, w_k)$: $x = w_0 + w_1 a_1 + w_2 a_2 + \cdots + w_k a_k$. The weights are calculated from the training data [197].

Any regression technique, whether linear or non-linear can be used for classification. The trick is to perform a regression for each class, setting the output equal to one for training instances that belong to the class and zero for those that not. The result is a linear expression for the class. Then, given a test example of unknown class, calculate the value of each linear expression and choose the one that is largest [197]. A regression approach can give smooth probability functions [212]. Figure 2.22 illustrates a linear classification equation (line in green) and the dashed green lines show the contours for p=0.25 and p=0.75, to be used for classifying two classes of elements (illustrated by the black and red points). From such illustration it could be seen that such equation is able to separate quite well the two classes elements in two distinct zones, consequently the linear assumption is good.

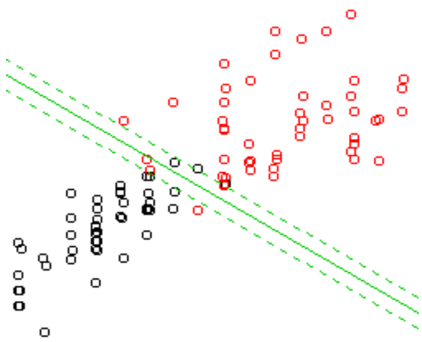


Figure 2.22 – Regression learning - linear classification equation example [212]

One example of the use of Numerical prediction is for the weather forecasts, which uses current weather conditions as input into mathematical models of the atmosphere to predict the weather. These are based on a set of equations (weather models) that interpret the physical laws that describe the hydrodynamic behaviour of the atmosphere. Figure 2.23 presents for instance a weather scheme used for forecasting calculation. The numerical weather prediction models are used for producing a weather forecast, which output from the model is studied by a forecaster before issued. This human-machine partnership is very important in producing accurate weather forecasts [213].

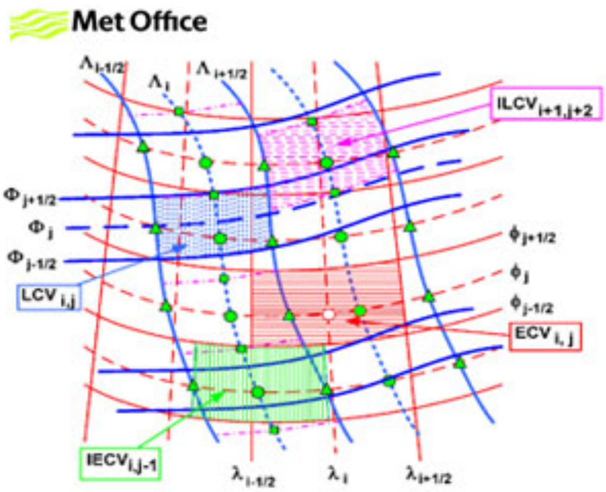


Figure 2.23 – A specific model scheme used in numerical weather prediction [214]

Data Mining

Data mining is the process of discovering meaningful correlations, patterns and trends by sifting through large amounts of data stored in repositories. The process must be automatic or (more usually) semiautomatic [197]. Data mining employs pattern recognition technologies, as well as statistical and mathematical techniques [215]. Such patterns are called structural because they capture the decision structure in an explicit way. In other words, they help to explain something about the data [216]. Data mining is frequently used to gain knowledge, not just predictions, and its discovered patterns must be meaningful in that they lead to some advantage, usually an economic advantage [197]. That is why its applications are extremely vast in economics, statistics and forecasting.

In a kind of short conclusion, there is machine learning, which aggregates all algorithms that supports a learning process in computers. By other side there is data mining, which uses data (independently to its source) to search patterns that intends to result in new valuable knowledge for the system. Thus, data mining is a specific sub class of machine learning due to its ability on the support for learning new knowledge. Some of the presented learning algorithms in this sub-chapter are also data mining algorithms as Apriori and K-means.

2.4.4 KMa Conclusions

Author analysed how the brain works in relation to the semantics representation to use it as inspiration in a possible KMa architecture or prototype focused in semantic handling. Consequently the two theoretical positions regarding the neuroanatomical distribution of the cortical semantic network and schematic models presented by Patterson et al. in [154], appeared to be a good inspiration to organise ontologies. Moreover, such schematic models could be translated to an ontology architecture, like as example, where it was represented brain connections can be converted to semantic mappings.

Then, following the same idea, it is presented some artificial intelligent semantic relatedness techniques/methods able to facilitate semantic adaptation ability to its connected system, inspired on human learning. Consequently, ANN, FL and OL were presented. Considering that the main objective is to reach a method that could handle with conceptual semantics, ANN is not so much considered as a solution due to its more appropriate relation for discrete or real values handling. Thus, even not rejecting it totally, a short conclusion is that ANN can be considered as a support (or as inspiration) for specific tasks that could be needed in an overall learning system able to handle with concepts. A possible example could be for acting as, or supporting a specific machine learning algorithm.

Despite to the high capacity of handling with imprecise inputs (uncertainty), FL is more appropriate to control systems then for providing learning capabilities, thus a priori, it was not considered as a necessary method to contribute to a semantic-based learning system. However, the author foresees that in the future it could be used to handle specific semantic-based learning operations as in semantic equivalences establishment between concepts in case of uncertain decisions.

By the last, OL is an important technique since it is able to represent knowledge with learning capabilities based on the employment of statistical methods in ontologies. Thus, due to the possible use of statistics in OL, an analysis of the MLN and BNs probability models was made. A MLN is similar to a BN in its representation of dependencies. For MLN nothing is invalid, just less probable. It can represent certain dependencies that a BN cannot (such as cyclic dependencies); on the other hand, it can't represent certain dependencies that a BN can (such as induced dependencies) [217]. From this, it can be concluded that both models can be applied, depending on the characteristics of the ontology to be modelled, or in an appropriate preparation of the data to be modelled.

These probability models still need support of learning algorithms to conduct the learning process as in readjusting their associated statistics. Consequently a machine learning literature review was made to help define these possible supporting algorithms. Such definition it was made categorised by the common classes of learning techniques types: Classification; Association; Clustering; and Numeric Prediction. Author concluded that all of these learning types should be considered to implement a robust and complete intelligent system able to maintain by itself its knowledge. However, due to industrial/business ontologies work mainly with products instances author found more appropriate to be focused on the instance based learning approach. These conducted author to the possibility of build an (inspired human) prototype where semantics are provided by domain experts but which could be updated (maintained) by external users in a similar way as humans do when they learn through the others. Humans build their own KB (lexicon) and update their ideas focused on what they learn from the interaction with others. They give more credibility to the people that they know that should know better a specific domain (e.g. professors). Thus, humans consider naturally such knowledge as right then they use for their lexicon establishment. And, by other side, other knowledge interactions with “not certified” people but mostly not less credible (e.g. colleagues) can be used to update their own knowledge. However this new knowledge will not erase previous ones until further confirmations for total rejections.

KMa is then proposed to be ruled by the analysis of the users’ interactions feedback, which works as the main trigger to the learning process that its related knowledge-based system could have. In additional this enforces the **reinforcement learning** approach in the KMa process. The intention is to have an OL, which will learn from its user’s usability in order to constantly improve the semantic interoperability between systems. The user feedbacks could be also used to get some specific patterns to improve the business as in recommendation systems.

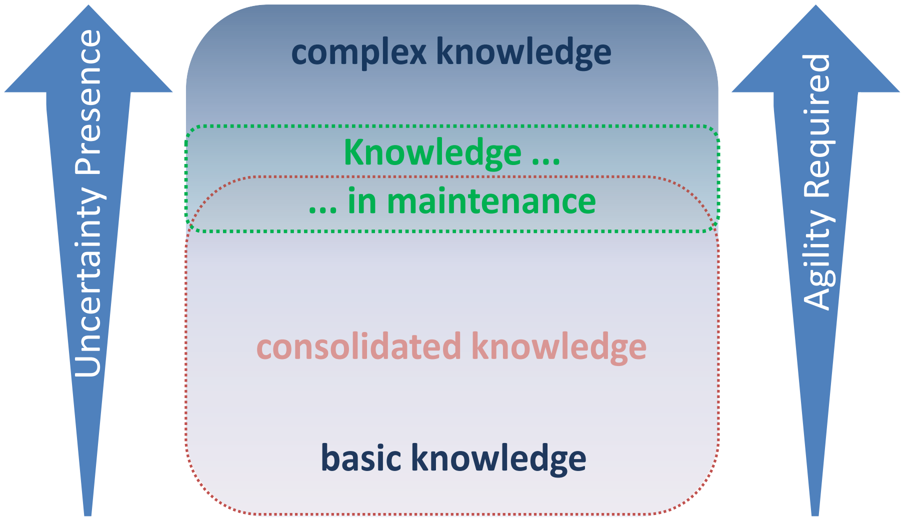


Figure 2.24 – Uncertainty and Agility presence in KMa

In additional to the presented conclusions, it is needed to emphasize that the insertion of statistics in ontologies leaded to the formation of the probabilistic ontology concept, which embodies the enabling of ontologies to represent uncertainty knowledge. Such uncertainty is present in knowledge proportionally to its complexity (as illustrated in Figure 2.24). How much complex the knowledge is how much uncertain it is. Knowledge in maintenance as presented in Figure 2.24, represents the knowledge that is in process of being consolidated by an individual or a system. Agility is other characteristic required to handle knowledge, which is proportionally required in comparing to the complexity of knowledge to be represented.

Matthews et al. in [143] suggest that agility is more than just the ability to respond to change; it is the ability to incorporate change as a way of managing the business [218]. In additional they state that how much an agile organisation is, the more dynamic and adaptive [219], better able to operate profitably in a competitive environment of continuous and unpredictable change [220][221], and more capable of gaining competitive advantage by intelligently, rapidly and proactively seizing opportunities and reacting to threats [222]. Agility is then considered as an enabler for more rapid maturity rate of knowledge change, creating greater capability for competitiveness. As a result, KMa should be performed through an agile system able to handle with uncertainty knowledge.

2.5 Knowledge Training (KT)

Knowledge training represents the needed training to deliver knowledge about something to a specific training audience. In this case, training is related to the knowledge needed to deliver skills to help develop competences on key people, which interact directly with the systems that semantically adapt to automatically retain interoperable. To reach this purpose it is needed to understand how to organise knowledge to transform it in appropriate and appellative training materials. Thus, it is needed to start by defining key training concepts to then introduce a possible model for training development.

2.5.1 **Training Key Concepts**

Learning is about to acquire new knowledge, behaviours, skills, values or preferences, and it may involve the processing of different types of information. Learning functions can be performed by different brain learning processes, which depend on the mental capacities of the learning subject/agent, the type of knowledge, which has to be acquitted, as well as on the socio-cognitive and environmental circumstances [223].

In other words, learning is an outcome, the end product of some process [224]. Such process can be categorised in two main categories: one more practical (or quantitative) category and other more complex one (or qualitative). The practical category aggregates mainly the learning as a quantitative increase in knowledge, which includes learning as memorising, storing information that can be reproduced; learning as acquiring facts, skills, and methods that can be retained and used as necessary. The complex category aggregates the personal aspect of learning, which includes learning as making sense or abstracting meaning, relating parts of the subject matter to each other and to the real world, and interpreting & understanding reality in a different way. Learning involves comprehending the world by reinterpreting knowledge [225].

From the previous learning definitions, it is possible to define and understand the objective of what **e-Learning** is. A usual accepted definition for e-Learning is: the delivery of a learning, training or education program by electronic means. E-learning involves the use of a computer or electronic device (e.g. a mobile phone) in some way to provide training, educational or learning material [226].

Training is a quite different from learning; it is the act of teaching someone how to do something, such as a job, or teaching the skills and attitudes that will have a direct impact on job performance, such as operations, human resources policies, or management and leadership [227].

Like in learning and e-Learning, the definition of **e-Training** is closely related to training. E-training is a term that is used to describe a number of techniques that use computer technology to deliver instructional material to trainees [228]. E-training is usually understood as e-Learning, although the core idea is the same, to deliver knowledge, but in training the knowledge is more task-oriented (e.g. how to do, how to use, how to install).

Overall, and with the increasing focus of e-Learning on knowledge-era workplace issues, and of e-Learning on real-world contexts, it is expected to see a convergence of models (e-Learning and e-Training) as they both attempt to achieve similar kinds of outcomes (e.g. higher-order thinking) but in different contexts and with differing degrees of generalization or specificity [229]. For now, e-Training is seen as a sub-category of e-Learning, a term mostly used for industry training.

Training delivery Types

A **web-based training** delivery method allows the course participants to extend the training time-span according to their needs and within a specific personal time allocation by accessing the materials via a browser or a network-connected application. By depending on general access to the course material, connectivity, performance and availability demands are also relevant issues to the training deployment.

With this delivery method each learner is able to execute the content in a preferred approach, e.g. always displaying examples of a certain learning topic first, as is the case of an example-oriented strategy. The learning objectives achieved can be validated by built-in assessments that allow electronic marks through the evaluation. Web-based training offers the highest flexibility, but also the necessity of additional discipline, because the user controls the organization of the learning process. Thus, the learner can be focused and concentrated to make a profitable use of learning materials [230].

Virtual classrooms are an alternative training delivery that has to be supported by a specific platform. It provides many of the advantages of the classroom training without the constraint of a fixed location. The participants join the course via web-based training or related technologies. This means that the participants are not required to come to a particular training location but must possess the necessary technical environment to participate. Since this delivery method requires fast Internet access it is only suitable for those persons who have access to such a connection. This means that it is more appropriate for smaller target groups and not for large-area dissemination. Virtual training has a higher demand on the performance and availability of the platform than the administrative use for classroom training [230].

Blended learning is another type of training delivery type characterized by a combination of all the other types (i.e. web-based with traditional forms, such as classroom training). A typical example of this would be a combination of technology-based materials and face-to-face sessions used together to deliver instruction [230].

Training Elements and Structure

In this chapter it will be described some key concepts about **training elements** or objects and structure.

A **training object** is any group of materials that is structured in a meaningful way and is tied to an education objective. In the creation of a training object, the author should consider how it relates to other existing training objects and other educational materials available in the platform [231].

A **training course** is an ordered process or succession of a number of lectures dealing with a subject. It is conceived in a way that meets the specific desires and expectations of a determined target audience. A training course is divided into several modules, according to the topics that are addressed. A **training module** is a small piece of a training course, essentially a lecture, with a very clear objective. Several modules of the same topic area can be grouped together to form a training course as illustrated in Figure 2.25.

Figure 2.25 – Training Course Structure

A **training programme** is a significant long-term training activity which comprises a set of training courses and/or training modules Figure 2.26 [232]. It's a construct conceived for training in specific skills focusing on a given target audience and using a selected delivery approach. Reference training programmes are those that are designed for reference target audiences (especially relevant within a given training environment) and that serve as orientation for targeted training execution [233].

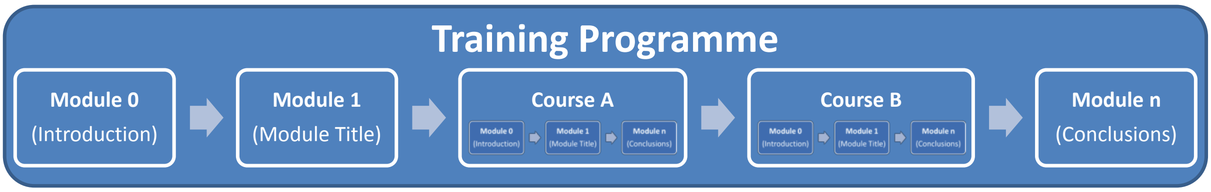


Figure 2.26 – Training Programme Structure Example

A **training curriculum** is the set of related instructional elements and content offers in a given field of study [234]. It's designed to establish the underpinning that is to be used to frame down training course elements. There could be several training curriculum areas, and each usually has at least one course defined with its direct contents of such area (Figure 2.27). Nevertheless there are some training topics that are used by other areas, so it is usual to find courses that exist in several training curriculums. Such relations give some complexity to the classification of these training elements [230].

A **dynamic training curriculum** is a curriculum, which its associated training courses are modularized in the sense to be able to pick up a module from them if appropriate, to a specific training programme. Figure 2.26 illustrates this example. The Module 1 could come from a course where all the other modules are not appropriate to be present on this specific training programme example. By this modularization it's enabled a dynamic curriculum. The dynamic training curriculum is therefore flexible, learner-centric and competency-based. This conceptual framework holds instructional elements (modules and materials) by focusing on atomic competences and skills within established domains [233].

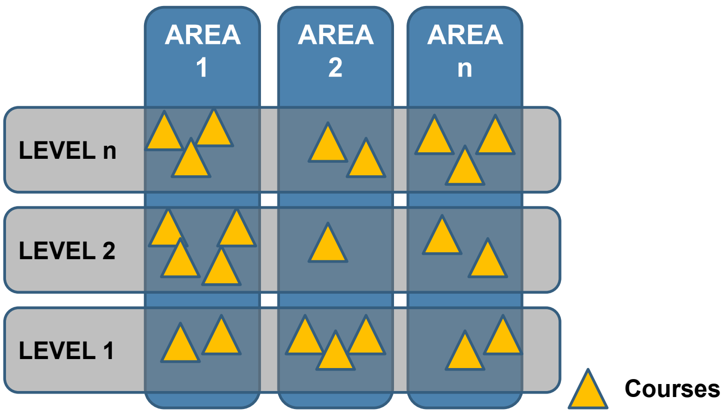


Figure 2.27 – Training Curriculum Structure Example

The main difference between a training curriculum and a training programme is that a training programme is designed to develop specific skills, and the training curriculum embraces all the fields of area of study. It is usual to find several training programmes within a particular training curriculum.

There are various ways to organize and deliver the training content, depending on the subject matter and the trainer preference. Figure 2.28 presents some structures types, which training lessons (courses or programmes) can follow.

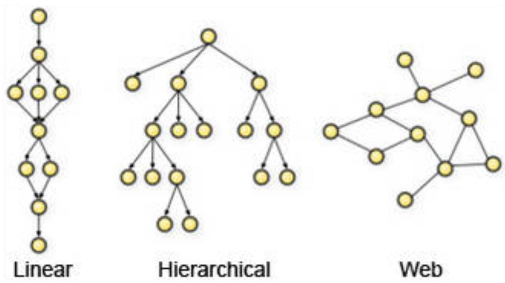


Figure 2.28 – Types of learning contents structures

The linear approach builds each lesson from a previous, in a single direction. The hierarchical structure organizes ideas with levels and branches, but still follows a top-down sequential grouping. The web structure is more flexible in terms of linking ideas, but is the most difficult to understand.

The training modules are organized and structured around the unit questions. Assessments inform the teacher and the learner about learner progress and at the same time, contributes to the learning process. Assessing student performance in an online course is similar to classroom assessment. Frequent assessment helps students keep pace with content requirements. A structure with assessments included is a good approach for producing an interactive course. These basic rules are often referred to as "Simple Sequencing". So, sequencing can be used to establish the pass/fail criteria, to determine a pathway through a number of Learning Objects (LOs), and to use pre- and post-assessment or testing to determine whether or not someone needs to engage with the learning content, and, finally, to determine the pathways for any remedial learning that may be needed by an individual learner [235] (see example at Figure 2.29).

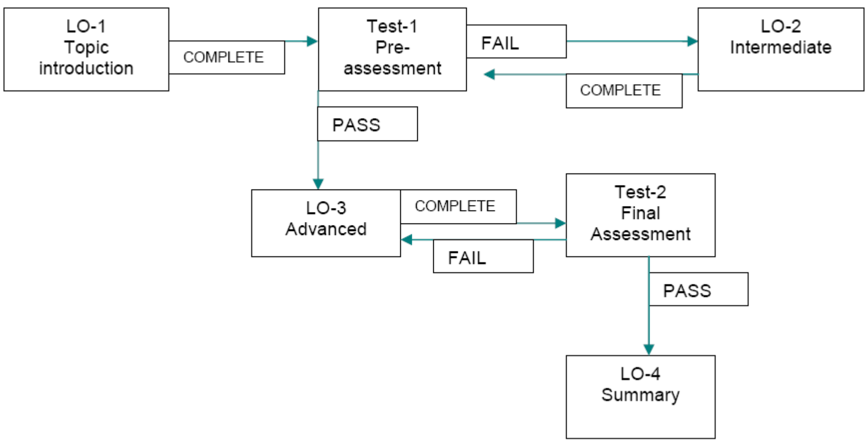


Figure 2.29 – A training structure with assessments enabling different sequencing possibilities

Metadata is the information about an object, be it physical or digital. As the number of training objects grows exponentially and the needs for learning expand equally dramatically, the lack of information or metadata about objects' places a critical and fundamental constraint on the ability to discover, manage, and use objects [236]. Thus, Metadata has a main role to the search engines within the learning platforms to locate and identify the LOs.

Specifications and standards in e-Learning enable different independent assets of learning to coexist for effective and better learning outcomes and also support properties like [237]: 1) Interoperability, making it possible to work with other LOs and with Learning Management Systems; 2) Reusability, allowing others the use of the objects created, even in different ways that firstly the object was designed for; 3) Accessibility, adding the information needed for quick and easy discovery so it can be found by other developer; and 4) Durability, by using the latest metadata standards so the lifespan long [231]. Many organizations like IMS[19], IEEE[20], ARIADNE[21], ADL[22], and AICC[23] are making standards in the field of e-Learning and most of the standards made by them are becoming the de facto standards in e-Learning [237]. These standards have been defined to structure training by providing metadata to represent its objects (e.g. multimedia content, instructional content, learning objectives, instructional software, persons, etc.). In additional, there are cases where a standard makes use of others. SCORM[24] is one of these examples. Its 2004 version introduces a complex idea called sequencing that is a set of rules that specify the order in which a learner may experience content objects, and which is an evolution of the underlying standards and specifications: 1) IEEE Data Model For Content Object Communications; 2) IEEE ECMAScript Application Programming Interface for Content to Runtime Services Communication; 3) IEEE Learning Object Metadata (LOM); 4) IEEE Extensible Markup Language (XML) Schema Binding for Learning Object metadata Data Model; 5) IMS Content Packaging; and 6) IMS Simple Sequencing.

A **Learning Management System** (LMS) is a software application for the administration, documentation, tracking, and reporting of training programs, classroom and online events, e-Learning programs, and training content. Some LMSs are web-based to facilitate access to learning content and administration. LMSs are used by regulated industries for compliance training. It is also used by educational institutions to enhance and support classroom teaching and offering courses to a larger population of learners across the globe [238]. Among many others, BlackBoard and MOODLE (Modular Object Oriented Developmental Learning Environment) are two well-known web-based

---

[19] IMS Global Learning Consortium - The mission of the IMS Global Learning Consortium is to advance technology that can affordably scale and improve educational participation and attainment. To ensure that the "Learning Impact" of technology-enabled innovation is achieved around the world (…) – Retrieved from the web at November 2011: http://www.imsglobal.org/aboutims.html.

[20] The IEEE Learning Technology Standards Committee (LTSC) is chartered by the IEEE Computer Society Standards Activity Board to develop internationally accredited technical standards, recommended practices, and guides for learning technology. – Retrieved from the web at November 2011: http://www.ieeeltsc.org:8080/Plone

[21] The ARIADNE Foundation is a not-for-profit association that aims to foster Share and Reuse of Learning Resources. – Retrieved from the web at November 2011: http://www.ariadne-eu.org

[22] The Advanced Distributed Learning (ADL) Initiative was established in 1997 to standardize and modernize training and education management and delivery and is part of the Department of Defense Office of the Deputy Assistant Secretary of Defense. – Retrieved from the web at November 2011: http://www.adlnet.gov/overview

[23] The Aviation Industry CBT Committee (AICC) is a nonprofit, membership-driven consortium dedicated to helping the training community get the most out of training technology. – Retrieved from the web at November 2011: http://www.aicc.org

[24] The Sharable Content Object Reference Model (SCORM) integrates a set of related technical standards, specifications, and guidelines designed to meet SCORM's high-level requirements—accessible, interoperable, durable, and reusable content and systems. – Retrieved from the web at November 2011: http://legacy.adlnet.gov/Technologies/scorm/default.aspx

LMSs, one being a commercial LMS and the other open source, respectively.

Blackboard born in 1998 at Cornell University and in the last years, it has grown exponentially acquiring several other companies, most notably, WebCT[25] (one of its rivals) in 2006. Blackboard offers numerous products, but their flagship one is the Blackboard Academic Suite, containing the content management system known as the Blackboard Learning System [239].

MOODLE is an open source course management system, also known as a LMS or a virtual learning environment [240]. MOODLE stands for modular object oriented developmental learning environment, and it is a well-known and widely used LMS [241]. It has become very popular among educators around the world as a tool for creating online dynamic web sites for their students [240].

2.5.2 A Methodology for Training Development

Methodologies establish the set of working methods that guide development processes, thus streamlining activities by providing the set of practices, procedures and rules that steers progress towards goals [233]. In this section, it is presented a literature review about relevant instructional models that could be used for supporting a training development methodology definition.

Instructional methodologies are founded on models for Instructional Design (ID). Such models provide procedural frameworks for the systematic production of instruction, and incorporate fundamental elements of the ID process including analysis of the intended audience or determining goals and objectives [242][243].

It could be stated that instructional models are guidelines or sets of strategies on which the approaches to teaching by instructors are based. Effective instructional models are based on learning theories. **Learning Theories** describe the ways that theorists believe people learn new ideas and concepts [243]. Theories about human learning can be grouped into four broad "perspectives": 1) Behaviourism - focus on observable behaviour; 2) Cognitive - learning as purely a mental/ neurological process; 3) Humanistic - emotions and affect play a role in learning; and 4) Social - humans learn best in-group activities [244].

There are other views about learning theories but all of them or at least in their majority have agreed on two main types: the behaviourism and cognitive approaches. Thus is mainly on these two approaches that author follows for training delivery.

Cognitive approach is more related to give <u>theoretical training</u> to the trainees. The various methods under this approach provide the rules for how to do something, written or verbal information, demonstrate relationships among concepts, etc. These methods are associated with changes in knowledge and attitude by stimulating learning. The various methods that come under cognitive approach are: lectures; demonstrations; discussions; computer based training; intelligent tutorial systems; programmed instruction; virtual reality [245].

[25] WebCT (Web Course Tools) was originally developed at the University of British Columbia by a faculty member in computer science at the beginning of 1996 – Retrieved from the web at November 2011: http://en.wikipedia.org/wiki/WebCT.

Behavioural approach is related to give practical training to the trainees. The various methods under behavioural approach allow the trainee to behaviour in a real fashion. The various methods that come under behavioural approach are: games and simulations; behaviour modelling; business games; case studies; equipment stimulators; in-basket technique; role plays [245].

Both previous presented training approaches have been used together for **competences** and **skills** development of industry people in European research projects (e.g. ATHENA; CoSpaces[26]). Cognitive approach has an important rule to acquire and consolidate new knowledge, and behavioural approach to put such new knowledge in practice, which is often made by using demonstration scenarios. Three different classes can define such skills development:

- **Skills – Know**. Are the skills to be acquired related to the knowing and understanding of something (theoretical knowledge of a field; the capacity to know and understand it);

- **Skills – Do**. Are the skills to be acquired related to the knowing how to act to certain situations (practical and operational application of knowledge to certain situations; be able to accomplish something);

- **Skills – Be**. Are the skills to be acquired related to the knowing how to be in determined environments (values such as an integral element of the way of perceiving and living with others and in a social context);

Thus, instructional training models should provide guidelines or frameworks to organize and structure the process of creating instructional activities able to provide skills to trainees to help them reach a pre-determined competence. Due to the intention of having the possibility of generating/adapting training contents to a specific trainee profile and/or objective, the author rather searched for Instructional Systems Design (ISD) models. ISD models are what DSchneider[27] calls ID methods, which is the systematic guidelines that instructional designers follow to create a workshop, a course, a curriculum, an instructional program, or a training session [246].

However, there are some researchers that distinguish ISD from ID models. Van Merriënboer in [247] states that ISD are used to guide the entire process of creating the learning platform, while ID modes are used in conjunction with ISD models to fill in the blank spots [248]. Clark in [248] also states that ISD models have a broad scope than ID, since an ISD model, typically divide the instruction design process into five stages (Analysis, Design, Development, Implementation and Evaluation – ADDIE), and ID models normally only focus on the design and somewhat on the analysis parts. From this last statement, it could be concluded that ADDIE is an ISD model. Even more, when Molenda in [249] states that ADDIE is virtually synonymous with ISD, and also refers that ADDIE represents the generic

[26] CoSpaces is an IP project funded by the EC under the IST Programme of the FP6, which overall objective is to develop organisational models and distributed technologies supporting innovative collaborative workspaces for individuals and project teams within distributed virtual manufacturing enterprises. Retrieved from the web at December 2011: www.cospaces.org.

[27] Daniel K. Schneider is senior lecturer and researcher at TECFA, a research and teaching unit in the faculty of psychology and education, University of Geneva. Retrieved from the web at December 2011: http://edutechwiki.unige.ch/en/User:DSchneider

ISD process. Thus, it could be said that ADDIE generally represents the ISD models type. Despite having available other ISD models like: Dick and Carey, Gerlach-Ely, Hannafin Peck, Jerrold Kemp, Knirk and Gustafson, and due to the general acceptance of the ADDIE model, as a simple but effective and systemic model compering to the others, author decided to follow it in his training organisation approach.

ADDIE

ADDIE is a widely used ISD model, which stands for Analysis, Design, Development, Implementation, and Evaluation. The acronym represents also the order in which is advisable to build the model, and each step has an outcome that feeds the subsequent step Figure 2.30.

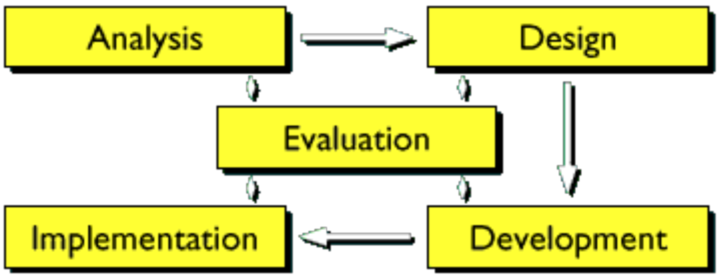


Figure 2.30 – ADDIE Instructional System Development Flowchart [250]

The Analysis stage determines training needs (e.g. analyse learner characteristics, task to be learned, etc.) and expresses them as information that is useful for training development. The ISD model requires that training fulfil specific needs. This is done through the generation and evaluation of such analysis elements as training objectives and target groups analysis. Analysis also considers the learning environment, any constraints, the delivery options, and the timeline for the project [251]. There are also defined performance measurement indicators for the tasks to be trained and an outline of the budget [250].

The Design stage is the ISD planning stage. Its purpose is to transform relevant content into concise, behavioural objectives, creating the instructional "blueprint" that will direct the development of all training materials, tests, and methods. Training requirements, target groups and outcomes identified during analysis are here mapped into goals and objectives, constituting the training courses, training curriculum and programmes [250].

The Development stage translates design specifications into training materials. Using the objectives, instructional approach, and input selections from design stage, the development activity produces instructional materials for both trainers and trainees, and evaluation instruments [250].

The Implementation stage focuses on details of training delivery/execution, as training of trainees and logistical arrangements. Work focuses on scheduling a training place, preparing an agenda, defining appropriate marketing, setting up the training environment, and delivery or distribution of instructional materials ensuring delivery of a training session able to captures trainees' interest. Moreover, and in order to reach a superior level on training materials, it is needed to train the authors on how they should develop the training objects in such a way that they exist in a format able to be deployed in an infrastructure which is capable to support various training forms and types [230].

The Evaluation stage ensures that training-under-development stays on track, safeguarding achievement of training goals and analysing system performance. A quality review process based on decisions and revisions for future course iterations can be made after evaluating the strengths and weaknesses in a completed training programme, thus ensuring achievement of desired goals. In each execution it is asked the trainees feedback concerning, e.g. materials quality, trainers performance, etc. This information is used for constant improvement of the training system [230].

2.5.3 **KT Conclusions**

Due to the constant evolution of systems, its interoperability is difficult to sustain. Thus, the knowledge of such systems actors i.e. people need to be updated with the changes of the systems. To help on this, it is needed to provide them training. However, if such changes are dynamically handled by the systems, training should be (if possible) also automatically updated to these changes and adapted to the systems actors' profiles.

To reach this purpose it is needed to understand how to organise training related knowledge and transform it in appropriate and appellative training objects. In additional, such related knowledge should be organised to assure training objects handling. Thus, the build of an ontology to represent training is an appropriate goal. Then, the presented training concepts definition can be used as a start point for it, where a thesaurus taxonomy of such concepts could be like the one represented by Figure 2.31.

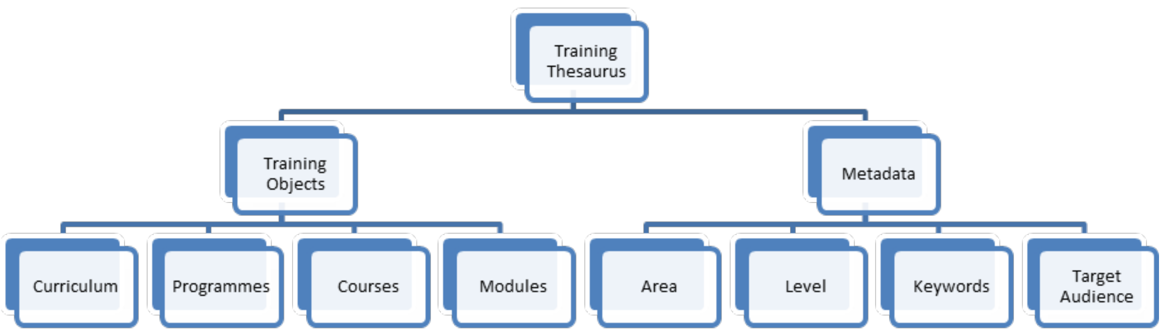


Figure 2.31 – An excerpt of a Training Thesaurus Taxonomy

From the presented thesaurus it is possible to build an ontology, following a methodology to be defined in the chapter 4. Such training ontology will be able to be used for generation of adapted training courses and programmes as wanted.

The author concludes that in accomplishment to this knowledge management approach, a methodology based on the ADDIE model can be used to efficiently organise the overall training development and its subsequent use.

3 Knowledge Based Framework for Dynamic Semantic Adaptability

This chapter intends to present the author's main conceptual contribution to this dissertation, which is a knowledge-based framework for dynamic semantic adaptability of industrial information systems.

3.1 The DIKUW Hierarchy

Due to the relevance of the knowledge use and interpretation in semantic adaptability of information systems, a careful background observation on its nature is required. As with any substantive concept, it emerges that knowledge is a multi-faceted and interconnected entity [252]. Thus, Bellinger et al. [48] defined knowledge through its related concepts, establishing a hierarchy between them. Data, Information, Knowledge and Wisdom (DIKW) compose this hierarchy, which "understanding" is achievied in the transition through the categories. However, the author, similarly to Ackoff categorisation of the human mind content [253], proposes the use of the "Understanding" as a new stage to the Belliger et al. DIKW hierarchy. With this new "Understanding" stage, it is established the DIKUW (Data, Information, Knowledge, Understanding and Wisdom) hierarchy. The reason is related to the level to which the information systems are actually able to represent. Today they are able to represent some knowledge but they are far way to reach the wisdom. Consequently, the author introduced the level understanding since it is a kind of middle stage between the knowledge and the wisdom level.

DIKUW hierarchy represents the knowledge creation and representation chain from the simple data to wisdom providing a foundation for exploring its dimensions conceptualisation. Each of the chain concepts is described through the conceptualisation of the previous concept on the chain.

Data is raw. It simply exists, and has no significance beyond its existence, in and of itself. It can exist in any form, usable or not. It does not have meaning of itself [48][254]. *For instance, if we are not looking at any specific object, the affirmation, "That object is a regular chair", don't tell us anything.*

By other side, **Information** is data that has been given meaning, by the way of a relational connection. This "meaning" can be useful, but does not have to be. Information embodies the understanding of a relationship of some sort, possibly cause and effect [48]. *For instance, "If a person can sit on that object, then it is a chair", this affirmation gives us one idea. The next time we see a person sat, no matter where, we will predict that he, or she, is sat on a chair.*

Knowledge is the appropriate collection of information, such that its intent is to be useful [48]. *For instance, "If the object is not heavy; has legs; has a back rest; and a person could sit on it, making more or less 90 degrees between the legs and the body. That's a regular chair (Figure 3.1)." This is a collection of information that almost certainly is useful.* There are two kinds of knowledge: (1) Explicit knowledge is knowledge that has been, or can be, articulated, codified, and stored in certain media [255]; and (2) Tacit knowledge, is knowledge that people carry in their minds, which provides context for people, places, ideas, and experiences [256].

Figure 3.1 – Regular Chair[28]

Understanding is an interpolative and probabilistic process. It is cognitive and analytical. It is the process by which one can take knowledge, and synthesize new knowledge from the previously held knowledge. The difference between understanding and knowledge is the difference between "learning" and "memorizing" [48][257]. *From the previous example, we could say: "The object illustrated in* Figure 3.2 *seems to accommodate a person to sit on it, making more or less 90 degrees between the legs and the body. Then, it is a regular chair, even if we think it is a strange chair." This statement proves that we learned from the knowledge acquired earlier.*

Figure 3.2 – Strange Chair[29]

Wisdom, is an extrapolative and non-deterministic, non-probabilistic process. It beckons to give us understanding, about which there has previously been no understanding. It is the essence of philosophical probing. It asks questions to which there is no easily achievable answer, and in some cases, to which there can be no humanly known answer, period [48][253][257]. *For instance, if we look to the object illustrated in Figure 3.3, at the first glance we do not know if it is a chair or a planetary system. However, since we have been talking about chairs, we can start thinking that perhaps, it could be a regular chair. Why not? Well, it is not a usual format, so it couldn't be a regular chair. We could state that we are concluding that "regular" is equivalent to "usual", thus a regular chair is something usual, where a person could sit, making more or less 90 degrees. All these thoughts indicate a presence of certain wisdom. Once again, is the picture presented in Figure 3.3, a regular chair? What's the size of each of the balls? Could a person sit on the bottom ball? From these statements, we can conclude that we are in a presence of wisdom since we are making a kind of philosophical probe of the previous acquired knowledge. Thus, wisdom could change the way we understand something.*

[28] Image retrieved from the web at August 2010: http://brainaudit.com/blog/2008/08/understanding-patterns-how-your-brain-thinks/
[29] Image retrieved from the web at August 2010: http://www.bentfabrication.com/suitcase.html

Figure 3.3 – Chair or Planetary System? [30]

3.1.1 **The "Understanding" Phase**

Today's research community is battling with the best way to collect data and information, in order to build KBs with the focus in reaching the understandability. Understandability is the ability of the systems to understand what is happening around. Using the presented DIKUW hierarchy, it could be said that researchers are now focused on the Understanding phase. Understandability will allow a computer system to draw conclusions, from knowledge patterns represented in a machine-interpretable form.

Diamantini stated in 2009 [258] that the Science of Interoperability developed several theoretical models to formalize the concept of understanding. Most, if not all, of them rely on the notion of mapping. Mappings can link directly two units of information, or the mapping can be indirectly defined through a third element acting as a common reference model (like an ontology concept, a meta-model construct, or the element of a global schema) [101][258][259][260][261]. This affirmation emphasizes the need of mappings establishment between existent information units or in third elements defined with the purpose of being the common reference to the involved organisations. This explains the importance of the knowledge creation and specially its reuse by organisations/enterprises, through the use of appropriate mappings.

In the pursuit to organise the knowledge creation process in a company, Nonaka and Takeuchi defined a theory of organisational knowledge creation, which states that knowledge is created in a cyclical trajectory simultaneously between ontological and epistemological planes, with spiral progression defining the conversion and mobilisation of tacit knowledge (Figure 3.4) [37][252]. Knowledge is initially created by individuals and then following this specific trajectory or process, becomes organizational knowledge. The epistemological planes are categorised by two types: the tacit and explicit knowledge. These types are related to the knowledge represented in the mind of people/actors and codified in any kind of data format respectively. Thus, Nonaka and Takeuchi identified that in the knowledge creation process, the tacit knowledge can be articulated to an explicit format and then to tacit again in each cycle of a spiral to create and represent the knowledge of a wide group of people. In additional, for each repeated process/cycle enclosing more people, an increase of the amount of knowledge is articulated resulting in a bigger ontological representation dimension. How much people are involved, a wider KR dimension (epistemological and ontological) is reached. This process can be performed in the opposite direction from an inter-organisational to an individual view, resulting in each iteration small ontological and knowledge epistemological dimensions respectively.

[30] Image retrieved from the web at August 2010: http://www.cribcandy.com/list=strange-chairs/175a5310d9ffe39c1b05ae14906517cc&pageoffset=0

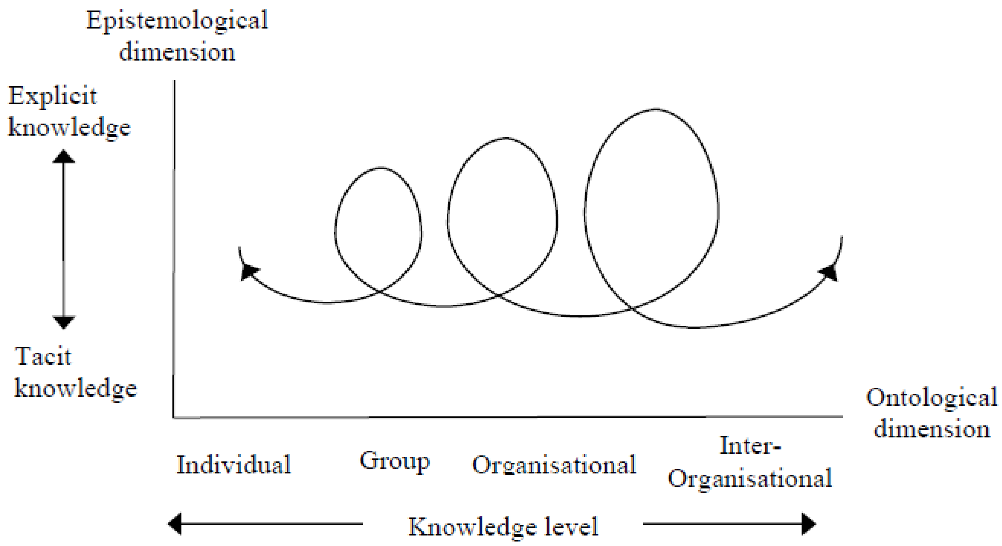


Figure 3.4 – Nonaka and Takeuchi's SOKC [37][252]

To reach understandability from the knowledge created, systems have to have intelligence able to manage knowledge and consequently to learn from it. The path is to define an overall system able to create and re-use knowledge, using mappings if needed, and following the Nonaka and Takeuchi's Spiral of Organisational Knowledge Creation (SOKC) (Figure 3.4). Such overall system is presented in the following two subsections, separated in two abstract (and small) ones, respectively related to the knowledge evolution from individual to inter-organisational view, and in the opposite way.

3.2 Semantic Interoperability Enhancer (SIE) system

The purpose of the SIE system is to contribute to make other systems interoperable by defining how semantic adaptability could be accomplished. The knowledge creation dynamics compliant with the semantic adaptability objective can be represented as a system in the sense that encloses a set of interrelated knowledge components working together towards the semantic interoperable system. Thus, it can be designed based on the Integration Definition for Function Modelling (IDEF0) standard structure. IDEF0 is a method intended to model the decisions, actions, and activities of an organization or system. This standard structure comprises a system based on inputs, outputs with feedback, controls and mechanisms of a determined function [262] (Figure 3.5). This system architecture is based on the training development management system proposed by Sarraipa et al. in [228].

The SIE system has as inputs, knowledge, tacit and explicit on a specific domain. It has mechanisms able to supply KREs to be used; and an entry control composed by the EI directives, which will conduct the system to the defined focus. Finally, it has the output that through appropriate knowledge reasoning is able to execute decisions. Such output functionalities communicate back to provide essential feedback on performance and quality of reasoning, thus enabling semantics adaptation towards excellence. These system elements are described in the following sub-sections.

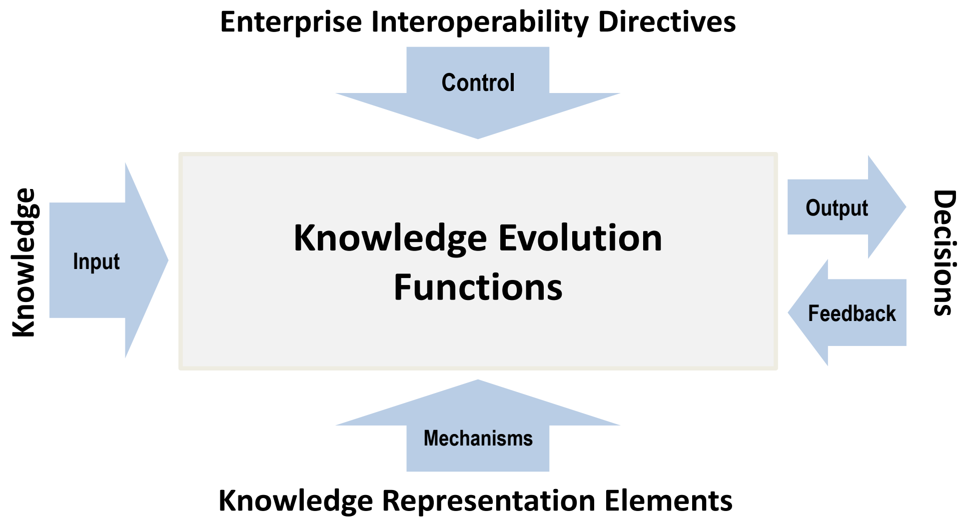


Figure 3.5 – SIE System

3.2.1 **EI Directives (Control)**

The control of the proposed system is ruled by a set of directives that are defined from the GC identified by the EI roadmap in [17]. Such directives encourage the way to which the control of the system should be lead. This is accomplished by bring directions/ideas identified by the EI roadmap and impose them to be used by the functions of the system. As an example: directive 1 denotes that the overall system to provide EI should act as a utility-like capability. Thus, this conducts to the implementation of interoperability functions (used by the proposed system) to be developed as a set of utility-like services. Therefore, the proposed system is synchronised with the overall EI research demands.

Directive 1 - Interoperability Service Utility

This directive denotes the overall system to provide EI as a utility-like capability. It pursues the creation of services to minimise the need and associated costs of enterprises, notably SMEs, to create information infrastructures that enable integration with different original equipment manufacturers in different sectors. It supports the use of basic information infrastructure that has information objects, ontologies, and metadata repositories as its core [17]. The ISU should address the following capabilities:

- Architectural design of the ISU as a “system of systems” to enable interoperability among peer systems without intervening in vital information exchange or exchange of collaboration knowledge, but sustaining interoperability through a mediation centre that will provide for coordination, negotiation and delegation functions.
- Semantic solutions for enabling the ISU operation and linking the ISU to the enterprises business context. New methods, ontologies (industry specific or generic) and tools for semantic mediation, service repositories, service matching, service retrieval and provisioning.
- Basic ISU services allowing end-to-end interoperability and value added ISU services

supporting co-creation of services and enabling innovation and value creation.

- Software and interoperability as service provision models of the ISU.
- ISU standardisation and standards-based operations.
- ISU business models and value assessment of ISU-related business models.
- Methods to link enterprise modelling to ISU engineering and tools to enable the transformation of enterprise models to executable logic descriptions.
- Accommodating cultural differences and multilingualism when needed

Directive 2 - Future Internet and Enterprise Systems

This directive is about researching new Web technologies for EI. It seeks to apply the concepts, technologies and solutions flowing from developments in Web technology to address the problems of EI. This will contribute to the advance of the implementation and uptake of Future Internet Services, establishing European-scale markets for smart infrastructure with integrated communications functionality [263]. Furthermore, new ideas would appear related to [17]:

- Model-based technological frameworks for achieving interoperability by design or through retrofit, during enterprise systems development and deployment, covering also widely-spread legacy applications, in an inter-enterprise or intra-enterprise environment.
- New examples of business-pervasive service platforms, specifically designed for SMEs or for vertical industrial sectors when needed, reaching new levels of adaptability, dynamic behaviour, and integration with the innovation process of the enterprise.
- New end-to-end examples of executable interoperability, in the form of coordinated intra and inter-enterprise workflow management, including technologies for automating semantic reconciliation and process flow management among diverse systems and services.
- New requirements for services and new service paradigms for the Future Internet, leading to new characteristics and properties for system openness, system adaptability and system integration in dynamic, programmable environments forming part of the Future Internet infrastructures.

Directive 3 - Knowledge-Oriented Collaboration and Semantic Interoperability

This directive supports the need for research in advanced technologies able to enable semantic interoperability and knowledge-based collaborations [17]. They could be new methodologies, techniques and tools for the discovery, capture and re-use of knowledge collaborative capabilities and services (e.g. folksonomies); next generation knowledge management systems, which are able to provide advanced knowledge services (identification, collection, representation, processing and exploitation) in support of knowledge-based collaborations; and ontology development and management approaches able to furnish the semantics for the Semantic Web and to committed practice communities.

Directive 4 - A Science Base for EI

This directive is focused on provide a science base for engineering EI, to support solutions of higher quality, dependability, and reliability. For this it relies on making EI more demonstrably cost-effective for end users; laying a long-term foundation for coherent and visionary EI research with broad impact; enabling the establishment of a multi-disciplinary EI research community; and by providing an infrastructure to support the diffusion of ideas, education, research, and training [17].

3.2.2 **Knowledge (Input)**

Traditionally knowledge and information are fragmented among many different places in organisations. For instance, knowledge such as best practice accounts, lessons learned and experiences about particular processes or procedures resides with the professionals, managers and engineers while, customer information, reports and procedures are often scattered across paper files and electronic databases. To leverage this scattered knowledge and increase organisational performance it is imperative to enable the flow of knowledge among and between individuals and groups within an organisation [264]. Knowledge acts as the input to the proposed system independently to its format (explicit or not). Once inside to such system knowledge has to be transferred or distributed among the organisation members, thereby promoting learning and producing new knowledge or understanding to enable semantic interoperability of systems.

3.2.3 **Decisions (Output/Feedback)**

An enterprise Knowledge-based system includes semantic features able to exchange decision-making information across its stakeholders. The knowledge from any domain can be modelled for decision making and applicable to building specific decision support tools [265].

The output of the proposed system will be used to facilitate reasoning from the users. Such output is directly connected to the system's KB that would be represented by ontologies. Thus it is able to support decisions or "answer" specific user's questions. It also acts as an interface to receive feedback from users. This feedback could be for instance obtained though the usage patterns of the ontology [266]. The prototype system has to track the ways users navigate for searching or reasoning the concepts and relations of the ontology. With an appropriate "ontology log analysis" it could be traced what areas of the ontology are often "used" and others which were not navigated.

3.2.4 **Knowledge Representation Elements (Mechanisms)**

A KRE is an element that facilitates the formal representation of the knowledge in a specific domain. Dictionary, glossary, taxonomy, thesaurus, ontology, and data representation standards are some of its examples. They are introduced here as technologic solutions (mechanisms) available able to represent knowledge and to be used by the proposed system.

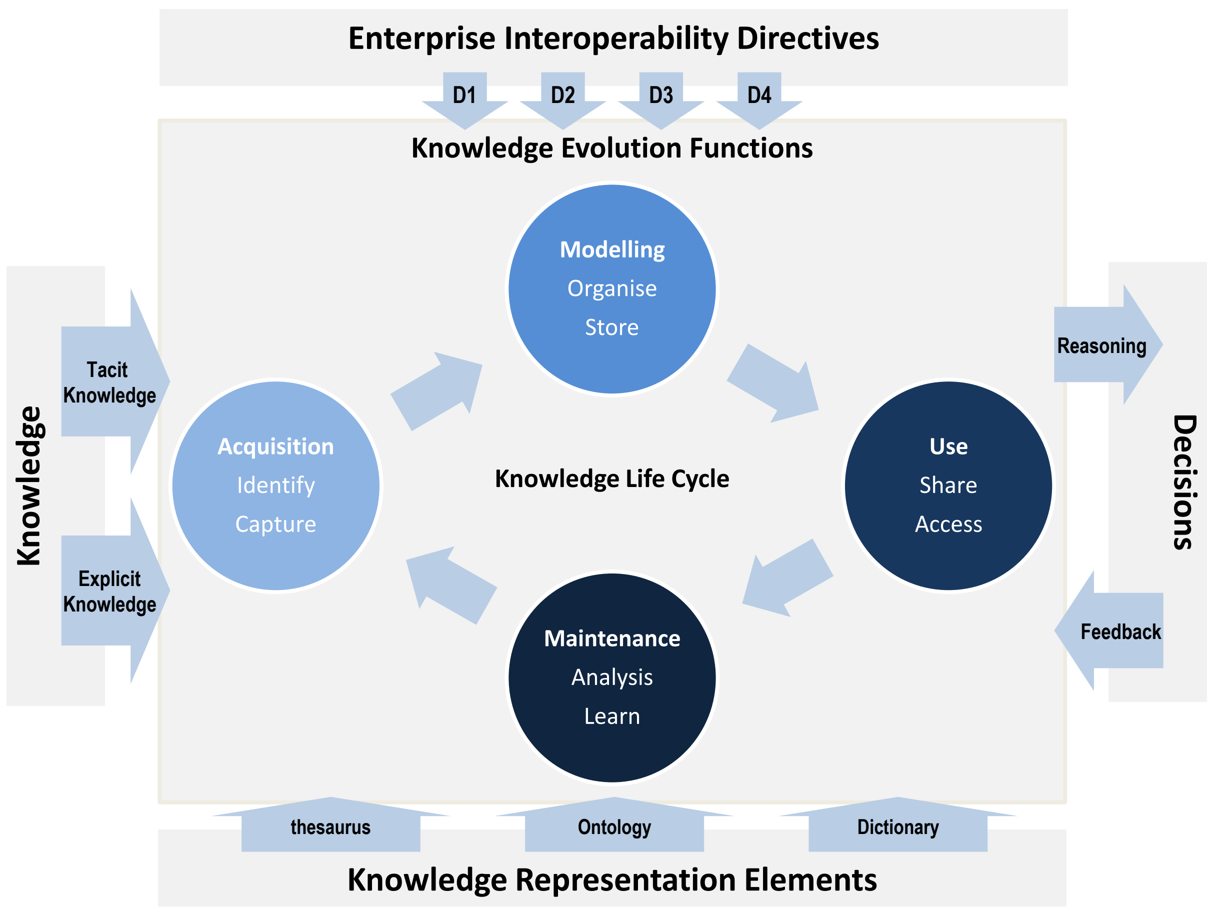


Figure 3.6 – Knowledge Evolution Functions of the SIE System

3.2.5 Knowledge Evolution Functions

The SIE system is rooted by a function based on the knowledge life cycle. A knowledge cycle is the process that knowledge passes through an organization as knowledge is identified, created, captured, shared, transferred, and utilized [267]. The author defined for such system a four stages' knowledge life cycle composed by: acquisition; modelling; use; and maintenance (Figure 3.6). Although, the function of the SIE system was correlated not only with the proposed four-stage knowledge life-cycle but also with the Nonaka and Takeuchi's SOKC. By this way, the function performs an entire knowledge life cycle, synchronised to the spiral conversion of knowledge cycle. The knowledge of the SIE system is acquired, modelled, used and improved on its way from a smaller view (individual or group) to a wider view (inter-organisational) (Figure 3.7). Thus, for each flow from identification to capture, store, share and maintain, knowledge is converted from tacit to explicit and then to tacit again.

Knowledge Acquisition

The knowledge acquisition is the process of knowledge identification and its capture to an explicit format to enable further digital use. Automated tools to knowledge identification and capture have been developed by the research community. However, how much informal the knowledge is represented on its source more its results decrease on efficiency. Moreover, if the process is focused on semantics of data even lighter solutions are available due to the complexity of the process of acquiring embedded meaning of a simple business domain concept. In additional this stage aggregates processes/methodologies able to acquire tacit knowledge to explicit through a human formalisation process.

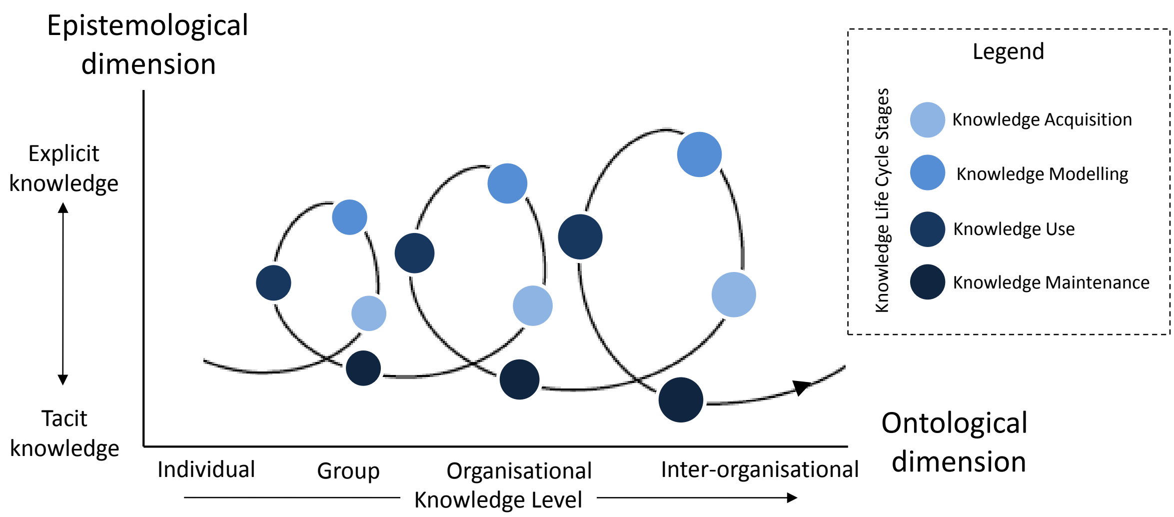


Figure 3.7 – Co-relation between the SIE knowledge life cycle stages and the Nonaka and Takeuchi's SOKC

Knowledge Modelling

The way the knowledge is stored and organised implies its further use. This stage receives the formal knowledge from the previous stage and organise it in a KB able to store and to enable further reasoning over the semantics of the concepts. The process, the KB architecture, the tools, and the formal languages used to model the knowledge are the main point of the KMo stage. The use of appropriate technologies as glossaries, thesaurus and ontologies should be considered depending on the purpose to which such knowledge will be used for.

Knowledge Use

This stage is centralised on the use of the stored knowledge to facilitate interoperability of the systems. Thus the way the knowledge could be shared and accessed by the various actors of a community is the main point of this stage. The formalism able to represent mappings between information models could facilitate the integration and use of various knowledge sources to the semantic adaptability of the information systems. This requires the ability of reasoning over the KBs to be the support services of the knowledge sharing. In additional, this stage is also related to use of the knowledge by reasoning functions to define conclusions or decisions and also for simple searching of products.

Knowledge Maintenance

The knowledge maintenance stage is focused on the improvement of the KB in order to be alive and updated accordingly to the knowledge evolution to which the system could be related to. Knowledge maintenance can be ruled, as instance, by the analysis of the users' interactions patterns (feedback), which works as the main trigger to the learning process that such knowledge-based system could have.

3.3 Enhanced Knowledge Training (EKT) system

Knowledge is always in evolution, thus a system (even technical) that handles it, has to be complemented with an effective knowledge transfer solution in order its actor be aware of all the changes. Levine and Gilbert in [268], stated that knowledge transfer can be accomplished through the support of incentives for generation of new ideas, specific structures to facilitate information discussion and flow within organisations, appropriate communication technology and effective training. The three first topics are also handled in the SIE system. EKT system works as a complement to the SIE system, consequently, author considered training as the main topic to be handled by the EKT system. By such both systems complementarity handling knowledge transfer, it is also accomplished the idea, also stated by Levine and Gilbert, that "knowledge transfer is only valuable when it is integrated into a set of policies for knowledge generation and capture" [268]. The policies mentioned here are related to the directives addressed in the control of knowledge acquisition and maintenance.

The purpose of the EKT system is to provide training programmes related to semantic interoperability systems solutions able to adapt automatically its contents accordingly to its target trainees' profiles and to the evolution of the systems to which it is related. For this purpose, author defined a system able to organise the training development in order to outline in what way the training objectives could be accomplished. It encloses a set of interrelated components working together towards the training process, defining each small step on the pursuit of training delivery objective. The identified training system follows the same approach as the one presented for the SIE system, being designed based in the IDEF0 standard structure (Figure 3.8).

Figure 3.8 – Training System Overview

The training system has as inputs, contents on the training objectives main themes. It has mechanisms able to supply resources for training materials development and delivery; and an entry control composed of the training principles, which will conduct the system to the defined focus. Finally, it has the output that in this case is the training delivery to the trainees. Trainees communicate back to system to provide essential feedback on performance and quality of training, thus enabling adaptation and change towards excellence. Such training system elements are described in the following sections.

3.3.1 **Training System Inputs**

The inputs of the training system are composed by the content sources. Two main sources of contents have been identified: internal and external. Internal contents come from own sources; the main purpose of having an identification of the internal sources is for internal governance of the contents used to produce training materials. External contents come from external sources to the involved organizations. The idea of having an identification of the external sources is mainly about intellectual property rights, and to facilitate trainees to follow to the training materials sources.

3.3.2 **Training System Control**

Training should follow an outcome-based approach where focus and drive are set on the specific outcomes of the training delivery towards a valuable and effective training experience. Clear objectives are defined, where a set of eight sound training principles are clearly identified to control the overall training system development and deployment (top of Figure 3.9).

1. The dynamic training curriculum principle defines that a curriculum is the set of related instructional elements and content offers in a given field of study [33][262]. A dynamic curriculum is a curriculum composed by training modules, which could be orchestrated to build adapted courses to specific characteristics, as target audience profiles and skills.

2. The reference training courses and programs are to be conceived in a way that meet the specific desires and expectations of a determined target audience.

3. Effective training implementation addresses how training execution should be carefully planned in order to generate the envisioned impact.

4. The methodology-based development principle establishes how the training development should be supported by a proper methodology in order to ensure quality management/assurance of materials accomplishing and its goals through a rational and logical path.

5. The valuable marketing and communication principle addresses how marketing and communication are important vehicles to reach targeted audiences and promote awareness of topics and value of the training services.

6. Only with appropriate technological infrastructures that will host and support training delivery is possible to realize the foreseen goals of the training services.

7. Accountable training activities are a principle that refers how an outcome-based approach to training focused on the results of delivery is supported by accountability.

8. And finally, Intellectual Property Rights (IPR) principle that addresses how the training

consortium should have a clear agreement on IPR for exploitation of developed training materials.

3.3.3 Training System Outputs

The training system outputs are based on two elements: the training execution and the training marketing. Training execution needs of an appropriated marketing to reach the target trainees. On the other hand the trainees are invited to provide a feedback in order to have a continuous training improvement.

Training execution ensures that the training is offered in a most flexible way to meet the different needs of trainees in order to achieve the desired results. The marketing or promotion essential goal is to create and sustain interest in training, and to promote the training services. The identification of the real interests of potential learners, ascertained through target audience analysis, is also vital for creating a product that meets the needs of customers. A focus on what the customer wants is essential to successful marketing. At the same time, this customer-orientation must also be balanced with the training objectives. Feedback mechanisms are intended to adapt, revise or re-plan the training execution in the various dimensions (curriculum, programmes, contents, etc.) and are mainly implemented through feedback questionnaires.

3.3.4 Training System Mechanisms

The training system mechanisms are assets that facilitate training execution. For each particular training course, these assets are assembled by the trainers themselves to provide training, e-Learning infrastructures as a vehicle to training delivery, and methodologies which provide directives for training development.

Trainers are expected to be able to properly identify and determine training requirements for each session [269]. Such work will require a specific design of the courses and programs – including evaluation schemes – according to each target group. The e-Learning infrastructure's main requisite is to make possible the delivery of virtual classroom and web-based training, and to give directions related to available traditional classrooms events.

3.3.5 Training System Functions

The training system presented is rooted in a well-established ISD approach, commonly referred as ADDIE [267]. It constitutes the EKT system functions (central area of Figure 3.9). The approach is especially relevant for the envisioned training system due to being simple, reliable, supporting self-adjustment and applicable to a broad range of training needs [233]. Such functions represent each step of the activities, actions, processes, and operations that embody the training system that should be used as a guide to the strategy and plan definition of the work needed to conceive, develop and implement the training activities required for an effective knowledge transfer.

The **Training Analysis (TA)** stage determines who, what, under which conditions, and how well the user must be trained. TA also intends to prepare tacit knowledge available at an organisational level to be transformed into appropriate explicit knowledge, in order the next target group (e.g. individual) could enrich with such knowledge transfer. Thus the identification of the training objectives and the specific target audience are the main purpose of this stage.

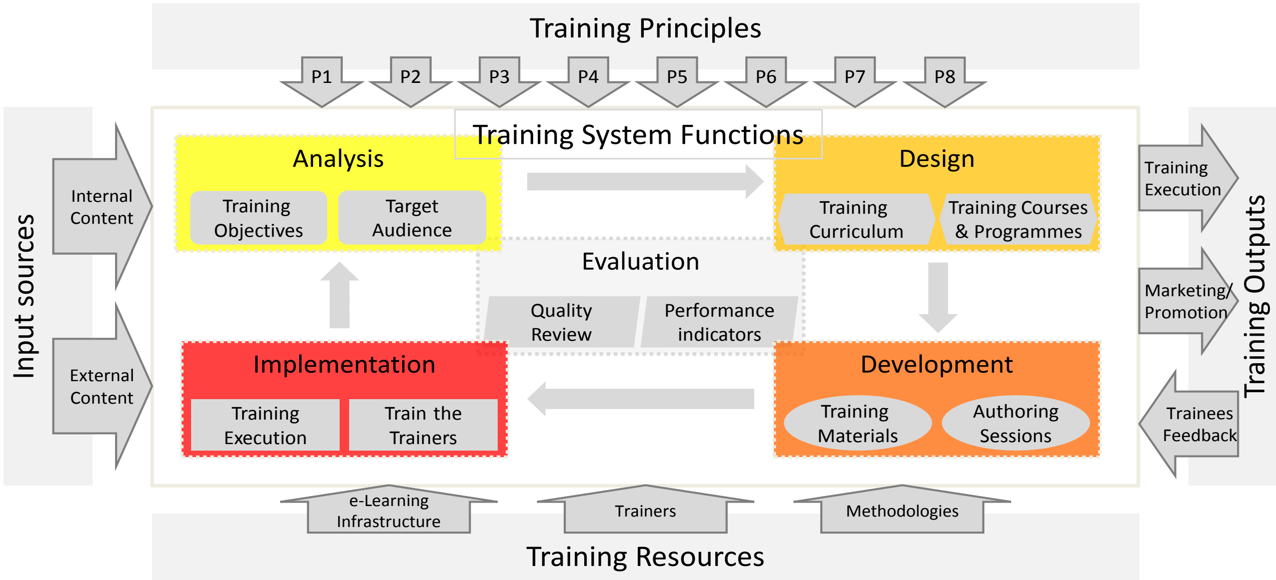


Figure 3.9 – Training System detailed view

The **Training Design (TDes)** purpose is to organise the knowledge to be trained (transferred) in appropriate modules, courses and programmes. Such elements should be designed for different levels of knowledge expertise (from beginner to expert) and different domain profiles, which could result in a training curriculum definition. This is accomplished by defining suitable mechanisms to integrate the training in the learning system, enabling its further customisation and training evolution control. The functional aim of this stage is to identify the main competency fields (or training domains) to be developed, in answer to the training needs identified during the TA stage.

The **Training Development (TDev)** stage transforms design specifications defined in the TDes stage into training materials. It makes use of Information Technologies (IT) environments specifications to develop training materials in appropriate formats/standards to facilitate its further host and transfer. The produced instructional materials are for both trainers and trainees, and evaluation instruments. This stage accomplishes the tacit to explicit knowledge transformation.

The **Training Implementation (TI)** stage focuses on to operate/deliver the training. It schedules the training service/place, and sets up the training environment instructional materials (and trainers) ensuring an adaptable training execution to the target audience. Through this action, explicit is transformed into tacit KR by the target trainees.

The **Training Evaluation (TE)** stage ensures the quality level of the training through trainees' feedback, which will be used for improvements development at the next TDev stage. Despite being possible to be executed at any time, TE stage is mainly fulfilled at the end of the training executions (TI stage).

The EKT system can be also correlated to the Nonaka and Takeuchi's SOKC as presented in the

Figure 3.10, where it could be verified that in each SOKC cycle is performed all the ADDIE model stages. Furthermore, each cycle represents a specific training implementation using as source a determined organisational knowledge to a small group of people with a specific target objectives and competences. This results in a knowledge transfer, where the amount of knowledge reaching the output target is smaller in dimension than its input.

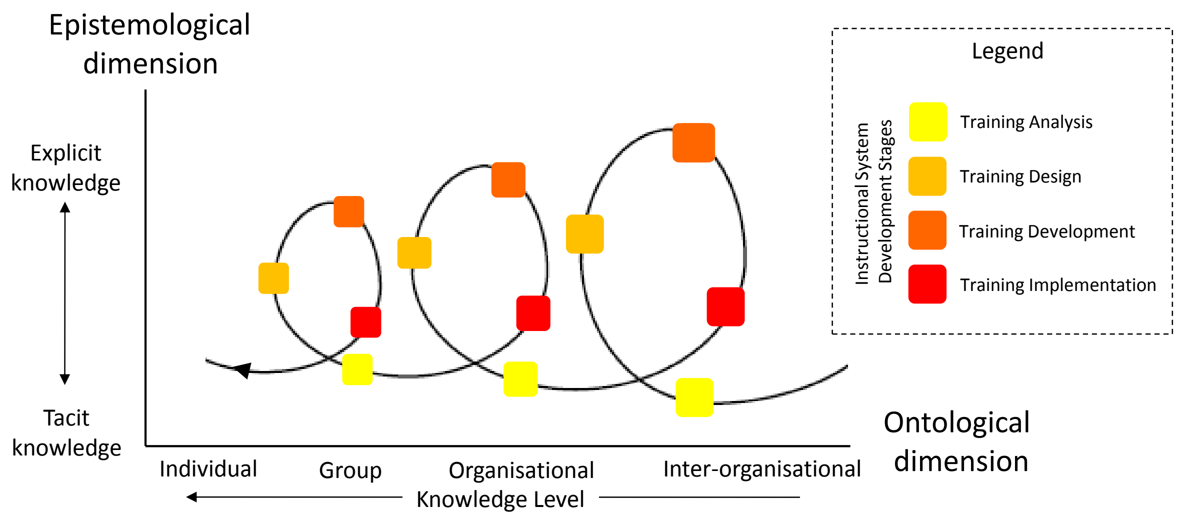


Figure 3.10 – Co-relation between the ADDIE model stages and the Nonaka and Takeuchi's SOKC

3.4 Dynamic Semantic Adaptability (DSA) framework assumptions

The proposed framework intends to represent a structural basis for supporting the development of systems for DSA resolution. In additional it provides a set of assumptions, to be addressed on such systems definitions, where its major guideline is to reach a dynamic adaptability of its handled semantics.

3.4.1 The SECI model

The introduction of the Nonaka and Takeuchi's theory in the definition of the framework systems is related to the intention of reaching the same dynamic behaviour handling semantics as such theory does with the knowledge creation. Knowledge creation is a spiralling process of interactions between explicit and tacit knowledge [270]. Consequently, it is needed to identify how these interactions lead to the creation of new knowledge, which should conduct to the spiral various phases or edges main characteristics identification and understanding.

The combination of the two knowledge categories, explicit and tacit, makes possible to conceptualise four conversions patterns [270]. These four knowledge conversion patterns include socialization (from tacit to tacit knowledge), externalization (from tacit to explicit knowledge), combination (from explicit to explicit knowledge), and internalization (from explicit to tacit knowledge) [271]. Figure 3.11 presents the identification of such knowledge conversion patterns and its relation/position to the SOKC, taking in consideration the direction to which the knowledge is evolving. The blue is related to the individual to inter-organisational direction, which is also related to the SIE system. The orange is related to the opposite direction, from an inter-organisational to an individual focus, which is also related to the EKT system.

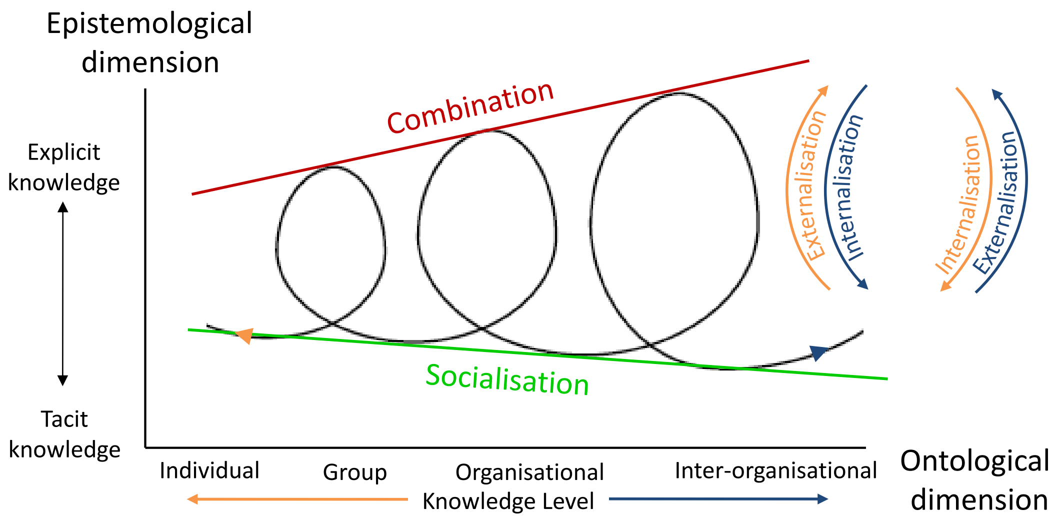


Figure 3.11 – The four knowledge conversion patterns of the Nonaka and Takeuchi's SOKC

Socialisation involves the sharing of tacit knowledge between individuals [270]. Such sharing is made through face-to-face communication or common experience [272]. In the KMa stage of the SIE system, there is socialisation related to the sharing of tacit knowledge from users (feedback) to the system administrators, which results in a discussion about the various users' inputs to decide what to maintain/update in the KR. In the KA stage of this same system, there are also some discussions that could still be part of this step (socialisation). Knowledge engineers and domain experts have discussions about concepts, definitions and procedures to agree on a common view in the domain. In the TA step of the EKT system, there is socialisation about the training objectives establishment and target audience identification.

Externalisation is the act of expressing/translating tacit knowledge into comprehensible forms that can be understood by others. An individual commits to the group and thus becomes one with the group [270]. In the final part of the KA stage all the acquired knowledge is transformed in a formal representation establishing by this way, in a shared common view of a domain. TDes step represents the externalisation in the EKT system by the fact that it defines training modules or courses to accomplish an overall training objective able to be represented by a curriculum.

Combination involves the conversion/organisation of explicit knowledge into more complex sets of explicit knowledge [270]. KMo is the SIE step where knowledge explicit is organised into more complex sets as from concepts to glossaries, thesaurus or ontologies. TDev is the EKT system "combination" step because it organises training modules or courses into "more complex sets" like specific adaptable training programmes.

Internalisation is closely linked to learning by doing; the explicit knowledge becomes part of the individuals' KB (e.g. mental model) and becomes an asset for their organisation [272]. In the KU step of the SIE system, users are able to search for specific products or services by specific reasoning procedures. They are able to learn from the stored knowledge in the system, becoming for instance more efficient in further utilisations. TI is the EKT system step that delivers training to people, which naturally accomplishes the knowledge transfer from an explicit into tacit format.

Some years after SOKC have been proposed, Nonaka and Konno in [270] (1998), despite to have focused their research in understanding the knowledge conversion from an individual to an entire organisation and vice-versa they presented their research focused in understanding what pushes or facilitates knowledge to be created as "new" or evolved dynamically in an enterprise. Thus, the (Nonaka and Takeuchi) knowledge creation theory was then focused around a new paradigm: the "Ba" characteristic. "Ba" is a Japanese word that refers to a common space or context, which is used as the foundation for the creation of individual and/or collective knowledge [273].

These changes in the Nonaka and Takeuchi knowledge creation theory focus supported the introduction of the **SECI** model. The name SECI comes from the first letters of the presented four knowledge conversion patterns (Socialisation, Externalisation, Combination, and Internalisation). Geytere in [272] stated that Nonaka and Takeuchi proposed SECI model for the knowledge creating process to understand the dynamic nature of knowledge creation, and to manage such a process effectively. It is consisted by three elements: the knowledge creation spiral; the knowledge assets and the "Ba" [272]. SECI model defines the knowledge creation as a spiral process, which results from the cyclical tacit and explicit knowledge conversions handled by the enterprises people (knowledge assets) in specific situations and in determined shared spaces or contexts – the "Ba". Figure 3.12 illustrates the SECI model and also presents each model phase adequacy to the presented SIE and EKT systems. Due to its nature, SIE is more focused in knowledge organisation, which conducts it to a higher adequacy to externalisation. By other side, EKT is more related to the knowledge transfer (to people) that makes it more adequate to internalisation.

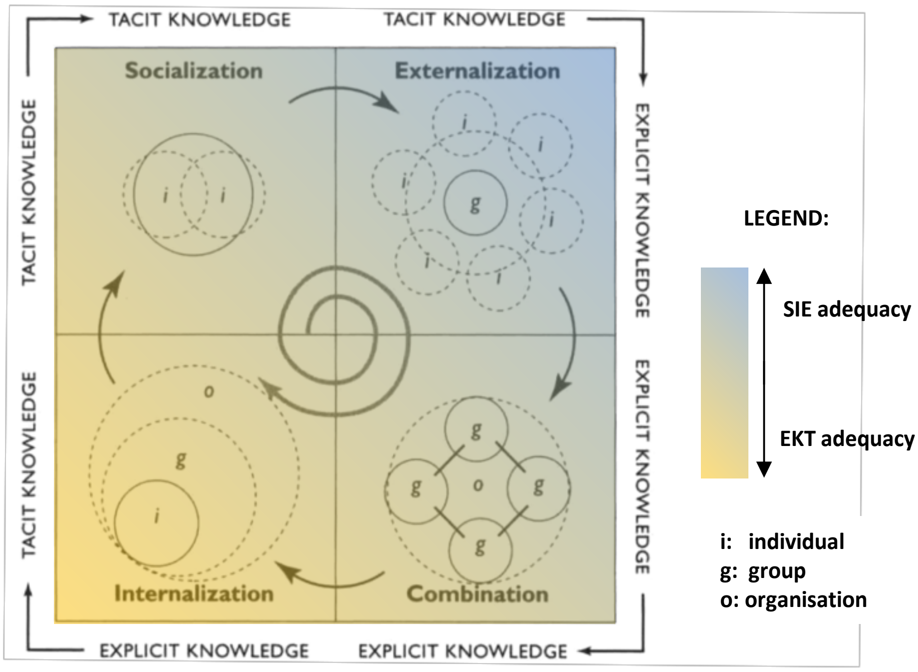


Figure 3.12 – SECI model and its adequacy to SIE and EKT systems (based on [270])

The main difference of the spirals represented by the SOKC (Figure 3.11) and SECI (Figure 3.12) respectively is that the first represent the ontological dimension of the knowledge created or converted between an individual to his overall organisation and vice-versa. The second is related to the new knowledge that the organisation knowledge assets are able to generate, when sharing the "Ba". For each new knowledge portion generated, an entire cycle of the SECI model is accomplished.

3.4.2 **Epistemological Dimensions of Knowledge**

Assudani in [274] presented two epistemological dimensions of knowledge: one related to the possession and other to the action/process. Epistemology of knowledge possession relates to the knowledge of something that is possessed by an enterprise. This possessed knowledge is demonstrated by enterprise's resources/assets or, as "knowledge of" the enterprise; and by new knowledge created and possessed after learning or innovation, or in other words, as "knowledge from" the enterprise. The Epistemology of action presents the knowledge as a process that mediate the relationship between the previous presented two forms, which determines if and how "knowledge of" is harnessed to generate "knowledge from" [274].

The "new" tacit and explicit knowledge created (just after one spiral cycle of the SECI model) can be exemplified to an enterprise as learning and innovation in the form of ''knowledge from'' the firm, and is exemplified as learning and innovation. This ''knowledge from'' feeds back into ''knowledge of'' the firm to become the new KB of the firm [274].

This new knowledge leads directly to the third key dimension of knowledge management, which focuses attention on the thought conditions that allow processes and tacit knowledge to evolve in the first place [275]. Scharmer's in [275] distinguished "Self Transcending Knowledge" as the third key dimension of knowledge, which has come more simply to refer to this dimension as wisdom. This is the force that drives the knowledge creation spiral of the SECI model.

While the knowledge management discussion of the 1990s revolved around the interplay of two forms of knowledge: explicit and tacit. In the current decade, it has been revolved around the interplay of three forms of knowing: explicit, tacit, and self-transcending knowledge [275].

The example of a loaf of bread can be used to ground these knowledge dimensions distinctions. Certain kinds of information about bread--like its weight, price, and ingredients--are examples of explicit knowledge. The activities of baking and producing the bread are examples of tacit knowledge, and the knowledge that enables a particular baker to invent baking bread in the first place is an example of not-yet-embodied knowledge. This "self-transcending knowledge" is tacit knowledge prior to its embodiment in day-to-day practices [275].

The concept of self-transcending knowledge proposes a distinction between two types of tacit knowledge: tacit-embodied knowledge on the one hand and not-yet-embodied knowledge on the other hand. The distinction is relevant because each of the three forms of knowledge - explicit, tacit-embodied, and self-transcending - is based on different epistemological assumptions and requires a different type of knowledge environment and learning infrastructure [275]. However, the last referred knowledge form is the more advanced and difficult to reach. Moreover, in the business area, the differentiation among markets with decreasing, steady, and increasing returns suggests that, in order to successfully compete for increasing return markets, leaders need a new type of knowledge that allows them to sense, tune into and actualize emerging business opportunities that is, to tap into the sources of not-yet-embodied knowledge. [275]

In a similar way to Assudani in [274] and based on the various presented knowledge assumptions, author also defined two epistemological dimensions of knowledge both related to possession, but focused in two distinct purposes: the self-enhancing and the self-transcending knowledge. Self-enhancing knowledge is related to the knowledge acquired from others in a kind of self-serving purpose. Self-transcending knowledge is related to the capacity of generate new knowledge or tacit not-yet-embodied knowledge and its communication to others. It is around these two knowledge creation epistemological dimensions that the DSA framework is established.

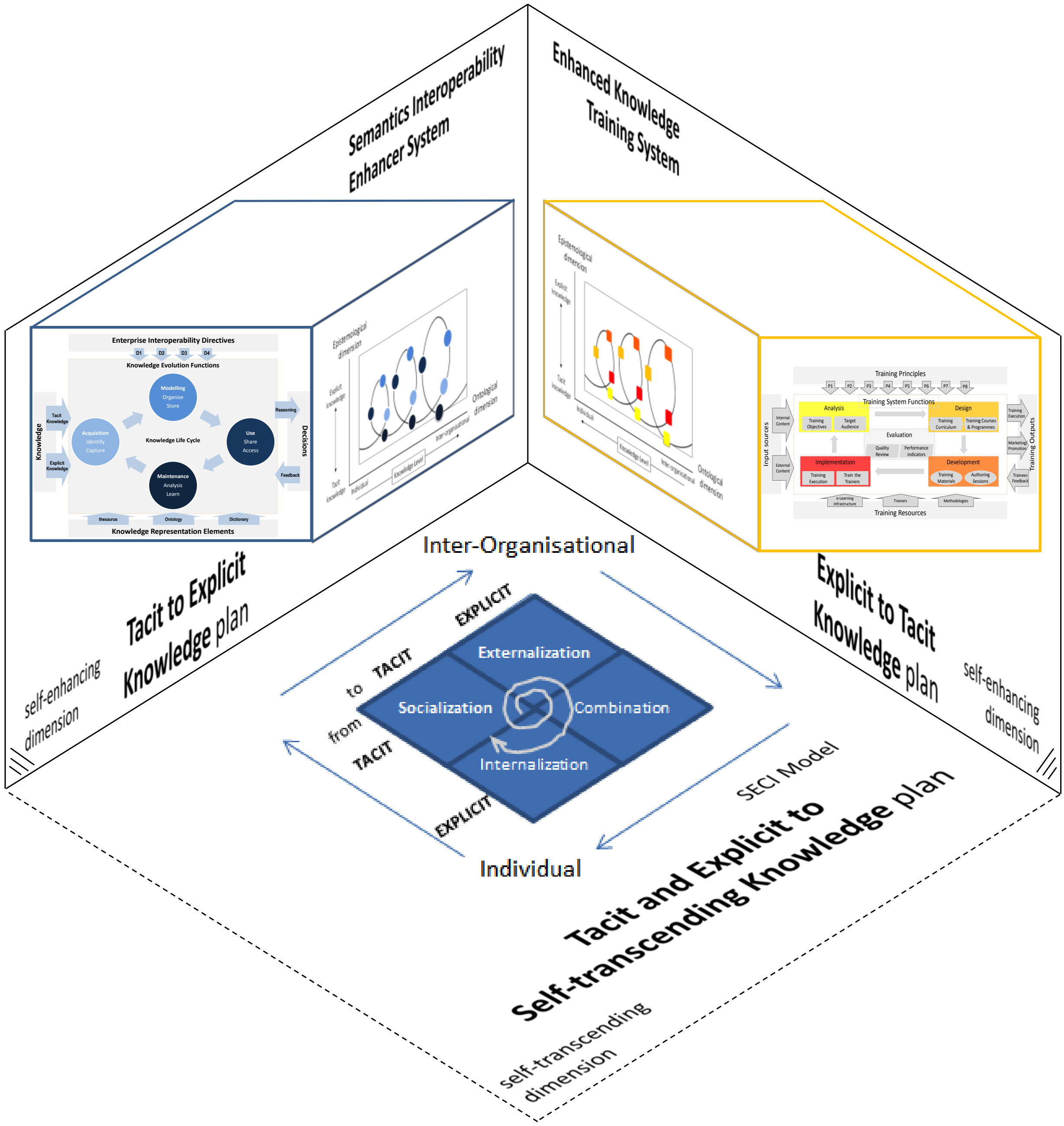


Figure 3.13 – DSA framework and its three knowledge creation plans

DSA framework accomplishes the SIE and EKT as two abstract systems. They act as a mirror since the first is responsible for knowledge organisation of the various tacit to explicit knowledge conversions resulted from an individual to an inter-organisational view transformation. By other side the second one is responsible for knowledge transfer from an inter-organisational to an individual view, through organised training executions, resulting in various explicit to tacit knowledge conversions, which is more or less the inversion of the first. Figure 3.13 illustrates these two abstract systems integrated with the SECI model representing three knowledge creation plans. SIE system works at the tacit to explicit knowledge plan and EKT works at the explicit to tacit knowledge plan, both handles knowledge in a self-enhancing dimension. The SECI model as stated by Nonaka et al. in [276], represents the continuous self-transcending process of knowledge creation (also mentioned as wisdom), thus acting in other dimension, the self-transcending dimension.

The proposed systems only handle the knowledge at self-enhancing dimension because as mentioned at beginning of this chapter, nowadays information systems are not able to reach the wisdom phase. They are (for now) focused in managing knowledge at most at understanding phase, thus the systems act in a self-enhancing purpose. They are able to learn from outside interactions but they are not able to really create new knowledge. Although, due to their relation and appropriateness to the SECI phases, as demonstrated in Figure 3.12, these systems should be also the basis for the self-transcending knowledge creation. A full machine solution handling at a wisdom level remains in the distant future. However, the framework oversees the introduction of human interactions to manage the self-transcending knowledge dimension. By specific knowledge maintenance procedures, such kind of knowledge could be integrated in the information systems.

3.5 DSA Framework Description and Conclusions

The proposed DSA framework builds upon the meaningful achievement of semantic interoperability between enterprise information systems, which must be achieved in a dynamic way. This framework provides a structure for supporting the development of systems for DSA resolutions. It follows the view that a system able to reach understandability from the knowledge created will have the capacity to learn from it and by consequence, will be able to manage its semantics in an adaptable way. Thus to reach such purpose the framework provides two abstract (SIE and EKT) systems to be followed in the development of such objective solutions.

Figure 3.14 presents the DSA framework with the SIE and EKT systems as two abstract systems, whose together handle the various phases of the knowledge lifecycle with the purpose to facilitate semantics adaptation of its connected enterprise information systems. In additional it is illustrated the framework main connections, which its systems could have. Enterprise information systems are by themselves one of these connections. They represent the communication networks that intend to receive support on semantic adaptations. Other shown connection type is related to the communications between administrators, knowledge engineers, domain experts and the SIE system. These people will have the role of organising and providing the main part of tacit knowledge to be used on the knowledge creation process. The connection to users represents other human interactions to enterprises outside people, like customers. Such interactions could result in specific feedbacks able to be used in knowledge updating processes. In the other side of the framework there are the system actors' interactions, whose represent the knowledge transfer from the systems to the internal enterprises people. Through this action, it is facilitated the creation of self-transcending knowledge, which could accomplish the SECI model cycle. These enterprises people would provide important feedbacks for the training (knowledge transfer) but also for the overall system knowledge improvement and update.

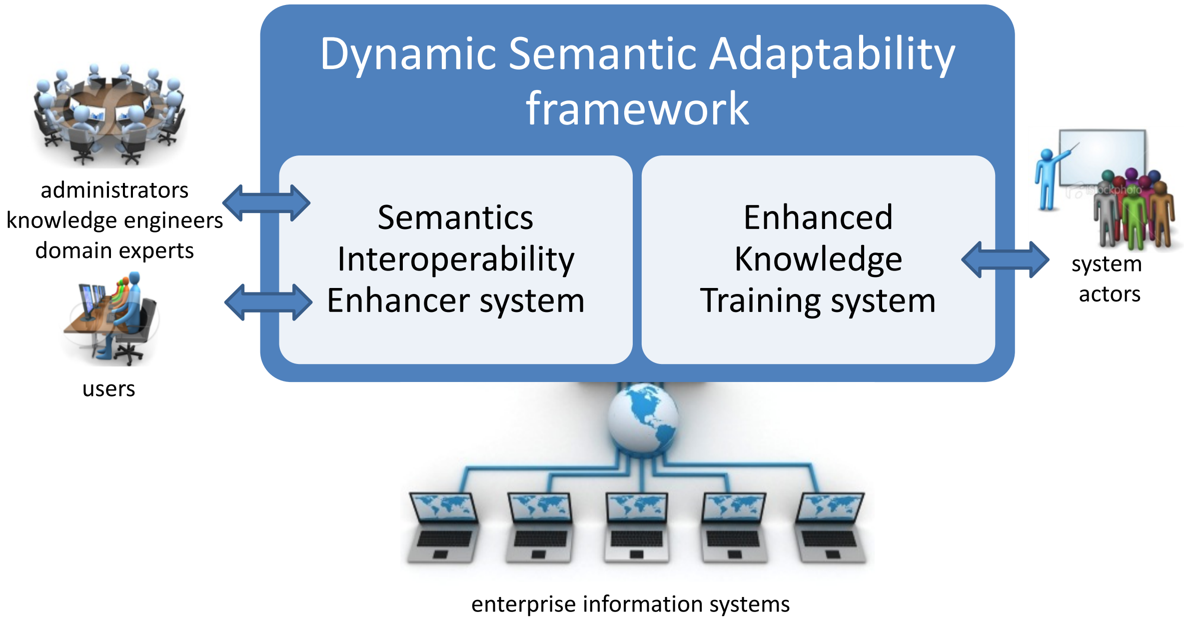


Figure 3.14 – DSA framework architecture

DSA framework provides the foundations to have a system to handle the semantics in parallel and close connection with the knowledge assets of the organisations. Thus, learning from them and helping them to learn and innovate, increasing their capabilities for sharing, transferring and creating knowledge to achieve organizational and business advantage. It provides a structure to be used as basis for the development of semantic interoperability resolution systems, taking in account two epistemological knowledge dimensions to which systems address knowledge management. About these dimensions, it assumes that due to the constant evolution of knowledge, systems are able to reach a dynamic semantic adaptation at self-enhancing knowledge dimension, but are far from reaching it at the self-transcending knowledge dimension. Thus, a complete fully machine approach reaching a dynamic semantic adaptation remains in the future. However, the way DSA framework is defined potentiates future systems overpass the understanding phase in the direction to reach wisdom.

The next four chapters of this dissertation intend to present and demonstrate the feasibility of the presented abstract systems, and to facilitate the approval of the approach presented on this framework, representing on that way the experiments conducted of this dissertation. Thus, each of these chapters provide solutions related to a specific part of the DSA framework.

4 Reference Ontology Building Methodology

A methodology is a comprehensive, integrated series of techniques or methods creating a general systems theory of how a class of thought-intensive work ought to be performed [277]. Methodologies formulate expertise (a.k.a. best practice) in performing a particular task in the form of guidelines that the reader can (re) use in performing a similar task [278]. There are methodologies and/or tools focused on contributing to solve interoperability ontology issues, which use operations, such as: ontology mapping/matching; ontology alignment; ontology translation; ontology transformation; ontology merging/integrating; and ontology checking. However, there is not an available methodology that addresses all of the above operations with cooperative and re-engineering ontology building processes for the implementation of a community reference ontology, and which is able to record those operations in a way that facilitates further traceability, solving semantic interoperability and facilitating the appearance of intelligent systems. This need has been recognised by the EI Research Community as reported in the EI Research Roadmap, in the research strategic view for GC: Knowledge-Oriented Collaboration: 'Research tools and methodologies are needed for acquiring, retaining and accessing the expanding range of knowledge available within individual enterprises and in virtual organisations, to enhance efficiency and productivity in collaboration, maximising the benefit of both long- and short-term knowledge sharing' [16].

4.1 MENTOR methodology

MENTOR – methodology for enterprise reference ontology development is a methodology that helps an organisation to build and adapt a domain reference ontology [84], with the characteristics mentioned before. MENTOR is supported by an ontology system constructed as a hybrid system. Hybrid systems are a special class of KR systems, which are constituted by two or more subsystems dealing with distinct portions of a KB and specific reasoning procedures [279]. Its aim is to combine the knowledge described by different formalisms in a semantic interoperable way. Thus, a hybrid specialised ontology system needs to be able to interoperate with other proprietary ontologies. Maedche et al. in [280], pointed out three approaches for combining such distributed heterogeneous ontologies:

- Ontology merging: using mediators;
- Ontology inclusion: in which the source ontology is simply included within the target ontology; and
- Ontology mapping: in which a part of the source ontology is related to the target ontology's entities.

Considering the listed options above, the author have chosen the ontology merging and mapping as the approach more adequate to combine the engaged ontologies in the system, since the focus is maintaining enterprise ontologies and building a new one to be their reference in the domain (Figure 4.1). Therefore, the described system is designed to facilitate semantic bridges between all the ontologies, which are related to ontology interoperability operations. In addition, MENTOR intends to address the interoperability ontologies operations considered relevant are [100][101]: (1) ontology mapping/matching, i.e. for each entity (concept, relation, attribute, etc.) in one ontology, a corresponding entity is defined in the second ontology, with the same or the closest intended meaning; (2) ontology alignment, i.e. the process of bringing two or more ontologies into mutual agreement, making them consistent and coherent with one another [281]; (3) ontology translation, i.e. the process of changing an ontology representation language keeping its semantics unaffected; (4) ontology transformation, i.e. the process that consists of modifying the structure and/or the properties of an ontology leaving unaltered its semantics; (5) ontology merging/integrating, i.e. building a new ontology starting from two or more existing ontologies with overlapping parts (merging), or extending some of existent parts (integrating); (6) ontology checking, i.e. ontology information inconsistencies checking, this is commonly performed by reasoners or theorem provers [282]; (7) ontology evolution/versioning, i.e. ontology domain changes or adaptations to different tasks, over time.

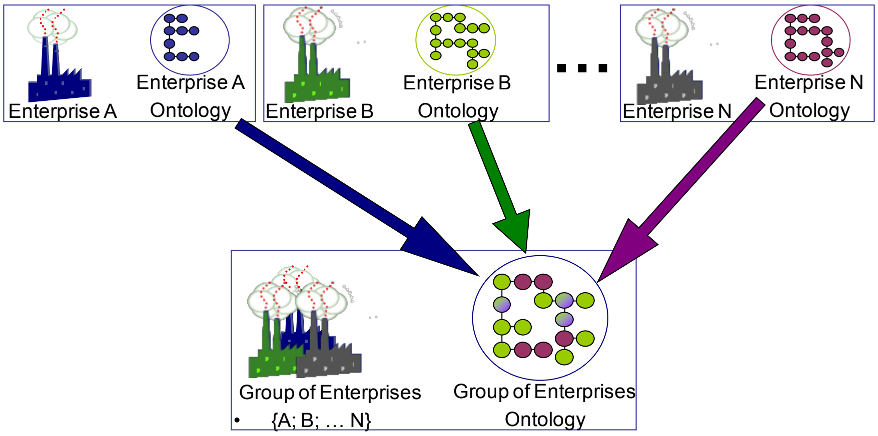


Figure 4.1 – Group of Enterprises Ontology – The Reference Ontology

However, before the MENTOR definition, the author made an analysis of a set of selected ontology building methodologies categorised by identified ontology building characteristics, i.e. ontology building from scratch, ontology re-engineering, cooperative building and merging methods (Table 2.2). In such analysis, it was identified that NeOn toolkit addresses all these mentioned characteristics, and by other side, Protégé tool, is the one that facilitates more the collaborative ontology building through its available plug-ins. However, NeOn and Protégé do not aggregate all the mentioned characteristics in an only one methodology, which could support enterprises to establish a common view on their semantics in a new ontology building process, at the same time as semantic mappings are established between their own KB elements and the new mentioned ontology [283].

MENTOR is proposed by the author in order to address all the referenced ontology building characteristics and interoperability operations, to facilitate the building of comprehensive capabilities for KR life cycle management able to provide semantic interoperability problem resolutions, in any information system domains. Thus, MENTOR provides a Mediator Ontology (MO) ready to record the information related to all the ontology operations with the objective of mediating the knowledge between the user and the KBs (ontologies) of the system.

MENTOR prototype (presented at sub-chapter 4.5) is mostly based on Protégé API (Application Programming Interface), thus it beneficiates from its main characteristics (e.g. language portability) and by its active community developments (plug-ins).

MENTOR provides several step methods, i.e. semantic comparisons, basic lexicon establishment, mappings and other operations among ontologies, among two distinct phases: the lexicon settlement (phase 1) and the reference ontology building (phase 2), with three steps each (*Figure 4.2*), whose are then composed of one or more actions.

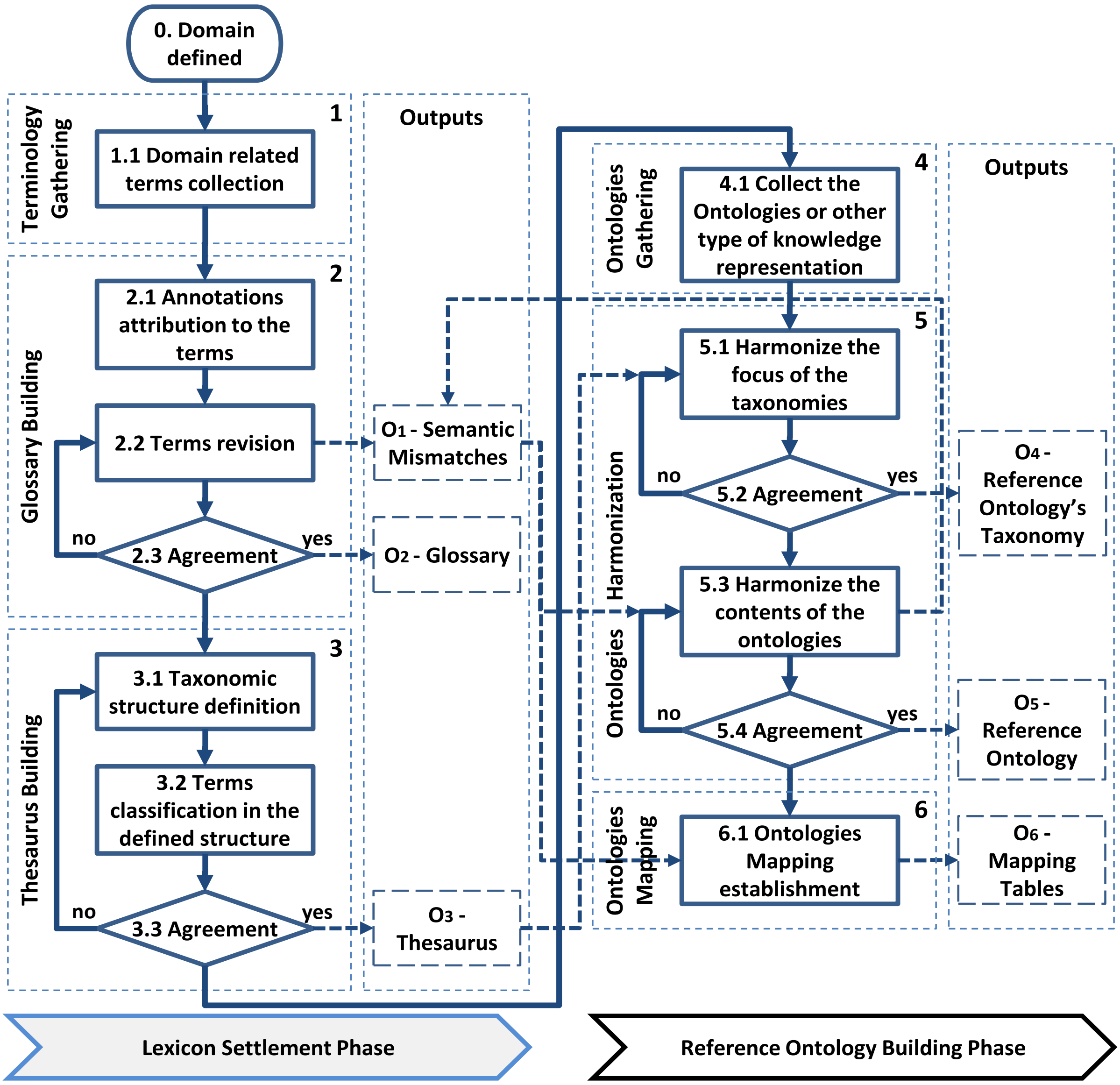


Figure 4.2 – MENTOR diagram

4.1.1 Lexicon settlement phase

The lexicon settlement phase intends to establish the vocabulary of a domain. Consequently, due to its capacity for representing domain reference concepts with associated meanings about a domain in an explicit way, thesaurus building is this phase main objective.

This lexicon settlement phase has three steps: (1) terminology gathering; (2) glossary building; and (3) thesaurus building. These steps were defined based on the UPON, which defines a set of workflows that also establishes a thesaurus for the domain before starting the ontology building.

The terminology gathering step (step 1) concerns the process of collecting all relevant terms or concepts in a specific domain previously defined. All the participants in the process should give their inputs. There is no rule from where the terms should come, since they are related to the domain established. Tools for automatic extraction of domain related terms can be used to accomplish this step.

Nevertheless, there is always need of a human checking before closing the terms list so as not to miss any domain terms. All the terms provided from the contributors are acceptable in this step, as nobody has authority at that moment to erase an other's participant term. Thus, the terms should be collected with reference to the contributor, in order to enable each contributor to provide the term's definitions during the next step.

Glossary is a specialised vocabulary with corresponding definitions. This vocabulary includes terms that are unique to a subject, and that have a special meaning in the field of interest. The definitions include descriptive comments and explanatory notes for the terms, such as synonyms, references, etc. A glossary is mainly used to unify knowledge sharing. The Glossary building step builds a glossary in the domain defined (Output O2). It starts with annotations attribution (action 2.1) to the terms collected in the step before. Then, each contributor should provide the annotations for his/her own terms. After having all the terms provided with annotations, it proceeds to the terms revision cycle (actions 2.2 and 2.3) to reach a reference definition. When the organisation members do not share the same natural language, this cycle adopts the use of a multi-language dictionary. Such a dictionary will help with the translations to the reference ontology language. The process for revision of terms can have four mismatches cases,

- Case 1: Existence of two syntactically different terms with the same meaning description – the action is to adopt one of the terms for being the reference in such semantics meaning. This process needs to be recorded as a semantic mismatch for future mappings (Output O1).
- Case 2: Existence of two syntactical equal terms with the same meaning description – the action is to erase one of them.
- Case 3: Existence of syntactically different terms with two different meaning descriptions – no action needed, both must be kept.
- Case 4: Existence of two syntactically equal terms with two different meaning descriptions – the action is to consolidate all the provided descriptions together in one of them and erase the other. In such a case, a new term could be proposed to the list if there is no agreement in the conjunction of the input descriptions and if the term to be born is not present in the terminology list. This process needs to be recorded as a semantic mismatch for future mappings (Output O1).

After a careful revision of all the terms with a successful agreement in their meaning consolidation, the glossary is defined from the terminology list in the domain specified.

The thesaurus building step is composed of a cycle where first the knowledge engineers define a taxonomic structure from the glossary terms (action 3.1), establishing some as thesaurus node terms. Afterwards, the other terms are classified into semantic proper paths in the existent taxonomic structure (action 3.2), down to the thesaurus leafs. If there is an agreement in the structure and in the terms classified, the thesaurus is defined (action 3.3). If not, the cycle starts again. The thesaurus defined (Output O3) will enhance the ontology harmonisation process in the next phase.

4.1.2 **Reference ontology building phase**

The reference ontology building phase (phase 2) is the phase where the reference ontology is built and the semantic mappings between the organisational ontologies and the reference one are established (right part of the Figure 1).

The first step comprehends ontologies gathering in the domain defined (action 4.1). Another type of KR could be used as input for the harmonisation ontologies process together with the thesaurus defined in the previous phase. The harmonisation method for building ontologies defined by an adaptation made from Noy and McGuinness in [284], and Jardim-Goncalves et al. in [285], proposes the development of a single harmonised ontology by two cycles, where first the structure is discussed until there is agreement on it (actions 5.1 and 5.2) and then the same process for the ontology contents definition (actions 5.3 and 5.4). The ontology contents definition process uses the semantic mismatches recorded in action 2.2 as an input in order to facilitate the harmonisation process. Nevertheless, in this process new semantic conflicts could be found. In such cases, new semantic mismatches should be recorded. With all the agreements accomplished in the ontologies harmonisation step, the reference ontology is finalised (Output O5). The next step is executed to establish the ontological relationships (mappings) between the reference ontology and each one of the individual ontologies (input ontologies identified in action 0). Such mappings uses the semantic mismatches recorded in the MO as an input to be considered, by the user on the mapping tables output accomplishment (Output O6).

Difficulties related to the natural language of the potential users of the harmonised ontology are likely to happen. To assist on it, the ontology building process could be complemented with a domain or multi-language dictionaries. However a special approach could be necessary. Chapter 0 presents one example of it.

4.2 MENTOR Enriched with QICM

All ontology building process needs a human intervention for at least to confirm a KR establishment. Consequently, a collaborative ontology building process, as MENTOR, should facilitates human discussions, for semantic and KR agreement. The employment of QICM in MENTOR, would effectively improve the approach to elicit knowledge from domain experts. Thus, this sub-chapter describes a suitable method to integrate for the human discussion phases present in MENTOR steps.

4.2.1 **Glossary Building using NG approach**

After gathering all the terms with proprietary (own enterprise) definitions, it is time to establish a group of reference terms with reference definitions – the glossary definition. The Glossary Building step is enriched with the NG approach and following the OntoWiki properties (Change Tracking, Commenting, Rating, Popularity, Activity/Provenance).

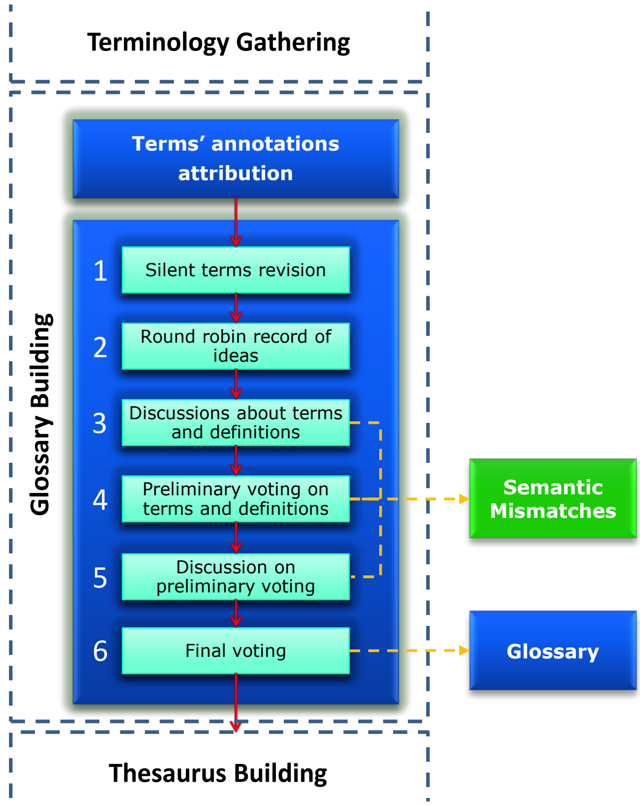


Figure 4.3 – Terms revision

This choice is based on the sequential nature of the adopted method, where a glossary building consists of a list of terms to be defined, aiming to discuss one by one from the beginning to the end without cycles. Since MENTOR prototype is planned to use web services, Metaplan technique is not the most appropriate, as it asks for a meeting for discussions, which make such implementation difficult to achieve. Surveys are also discharged because of its absence in discussions. Therefore, the glossary building is accomplished following the six NG's steps as depicted in Figure 4.3:

- Step 1: all participants review alone the overall list of terms rating and commenting them;
- Step 2: participants write down some ideas about each corresponding term and definition (equivalent terms identification);
- Step 3: the system manager starts and leads a group discussion about the terms and their definitions. The objective of this group discussion is finding which terms should compose the glossary being the reference ones and find a proper definition for each term;
- Step 4: a preliminary voting process is taken to see if there was an agreement between the participants. All members vote if they agree with each term and definition;
- Step 5: a new discussion takes place to clear up any doubts about some specific terms or definitions;

- Step 6: a final voting session defines the output of this step, the glossary. This glossary contains all terms and their definition decided by majority.

From the last steps of this approach it is established a list of the semantic mismatches records related to each term linked, between proprietary and reference ones.

4.2.2 **Thesaurus Building**

After having the glossary built, it is performed the thesaurus definition. In this process it is discussed the taxonomic structure of the thesaurus.

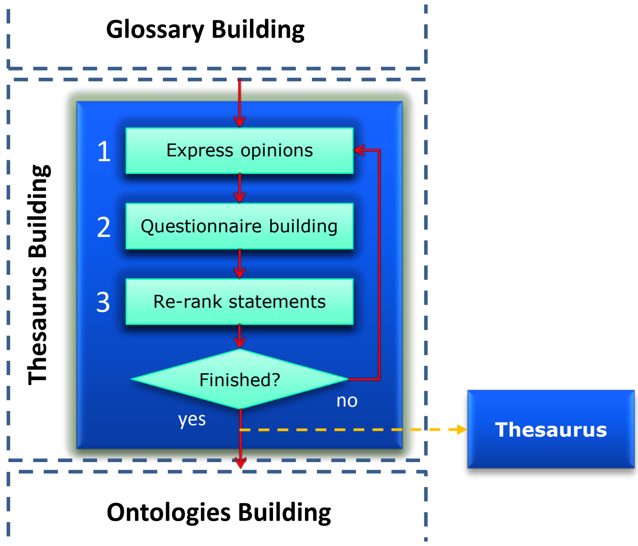


Figure 4.4 – Taxonomic structure definition and terms classification

Discussions around a taxonomic structure definition are more complex than for a simple glossary definition. It requires higher conceptual discussions then in the semantic attribution to the terms. For instance, it is needed to know the entire domain, to be able to relate and classify the existing terms in a common structure. Thus, as described in the literature review chapter, when the discussions are complex, Delphi is a good method to follow. Delphi facilitates a kind of re-rank answers enabling sometimes to reach a faster solution, avoiding repetition of extra discussions. Author proposes the Delphi method using 3 rounds, as the best solution for the thesaurus building purpose (Figure 4.4). First it is needed the definition of an expert panel, which should be composed by knowledge engineers and some domain experts. Then it follows 3 rounds in cycle if necessary:

- Round 1: Express opinions – Every member of the discussion group is asked to express any opinion about the structure of the thesaurus. This discussion is used to identify questions, for example, “which are the main classes able to be used to classify a specific business domain (section)?”

- Round 2: Members answer to a questionnaire built with the results of round 1 (this could be open to other domain experts from the involved companies);

- Round 3: Participants re-rank their initial statements with the results of the round 2.

These rounds are performed at least once for each taxonomic level definition and for each branch (Figure 4.5). It should starts by the first level classes identification from the glossary terms (Level 1 of Figure 4.5and then through each branch and level by level (Level 2 of Branch 1 and then Level 3 of Branch 1 before change to another branch - Figure 4.5).

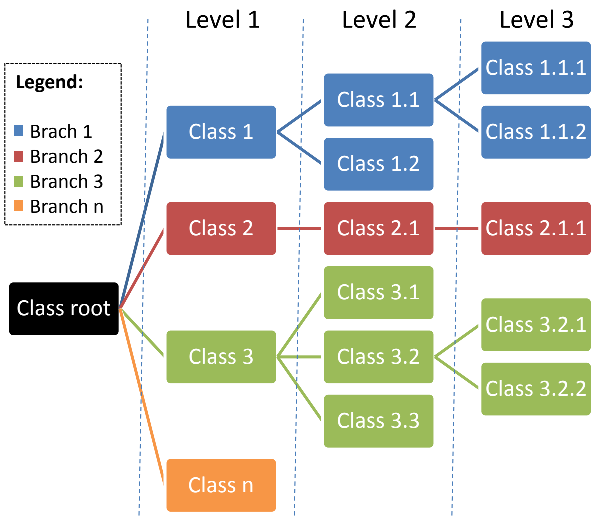


Figure 4.5 – Taxonomic levels

Once the thesaurus is defined it is time to proceed to the reference ontology building. It starts by gathering the various proprietary ontologies from the involved enterprises, plus the thesaurus defined previously. Then, the ontology's harmonization step starts, through two cycles: ontology's taxonomy harmonization cycle; and ontology's contents harmonization cycle.

4.2.3 **Ontology's Taxonomy Harmonization Cycle**

The ontology's taxonomy harmonization cycle has a similar objective as the thesaurus building step. Consequently, the method to follow is Delphi as illustrated in Figure 4.6. Although, in this case the Delphi rounds are applied to the whole structure and not level by level as in the thesaurus building step. The difference is justified by the fact that in this case there is a taxonomic structure already agreed by all that is the thesaurus taxonomy, which could work as a possible basis. Other difference could be related to the fact that thesaurus building uses the list of concepts identified for the glossary as a pre-identified input. And for the ontology it could be focused in other concepts not included in thesaurus due to its common sense or to its relation for specific purposes as for technical solutions used in the related domain or project information systems.

The ontology's taxonomy harmonization cycle comprehends an expert panel composed by an administrator (knowledge engineer) and domain experts of the involved enterprises' participants and through four Dephi rounds:

- Round 1: Express opinions – Every member of the discussion group is asked to express any opinion about the reference ontology taxonomy. For instance one usual question to be considered at the first cycle is: what are the main questions to which this ontology should be able to answer? This helps on taxonomy definition since such identified questions answers could contribute for the main classes' identification, also facilitating its ability, to represent things that could compose the mentioned answers. On this round another important issue is related to the re-use possibility of proprietary ontologies parts (e.g. properties, rules; entire branchs) or the thesaurus taxonomy, which were gathered on the MENTOR step before. Thus, it is around these issue that discussions able to facilitate the expression of opinions are supported, normally resulting in questions identification;

- Round 2: Members answer to the questionnaire built with the results of round 1;
- Round 3: Participants review all answers and consequently if needed the administrator update the reference ontology taxonomy.
- Round 4: In a similar way to the INTEROP collaborative approach to build a glossary (sub-chapter 2.2.3), participants present their opinion about the ontology taxonomy built by choosing one of the options: review (rev), reject (rej), accept (ok) or ignore (blank). If there is not a common agreement, they start a new cycle iteration. Otherwise the ontology taxonomy is defined.

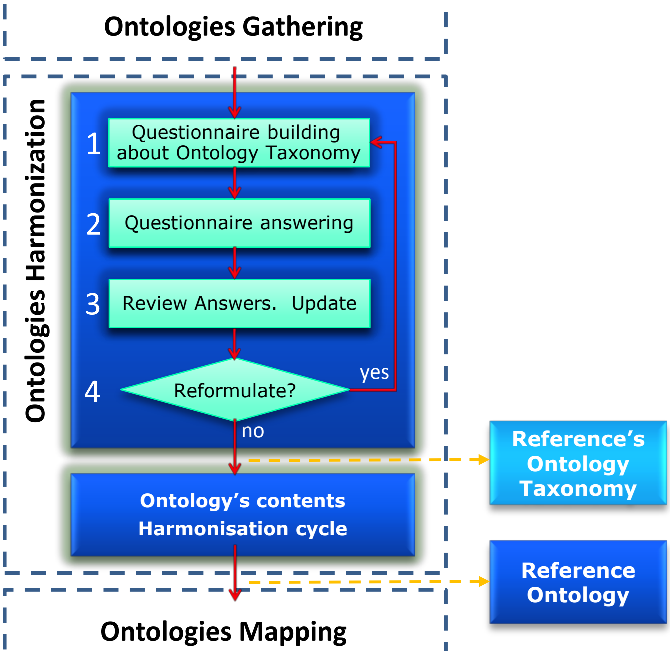


Figure 4.6 – Ontology's Taxonomy Harmonization Cycle

4.2.4 Harmonization Cycle for Ontology Content Focus

With the Reference Ontology Taxonomy defined, it is needed to establish its properties and rules. The process applied is similar to the previous one. Thus, it follows the Delphi method but by this time it is applied to the properties and rules (Figure 4.7), comprehending the same expert panel, with the same four rounds:

- Round 1: Express opinions – Every member of the discussion group is asked to express opinions about the properties and rules of the new ontology classes (identified in the step before). Such discussions have the objective of defining a set of questions (e. g. is the rule R2 able to regulate a specific object characteristic?).

- Round 2: Members answer to a questionnaire built on round 1.

- Round 3: Participants see all the answers and based on that they are able to review their initial answers defined at round 2.

- Round 4: If needed the administrator reformulate the reference ontology properties and rules. Due to this study of the ontology classes' characteristics, could appear new considerations that would conduct to changes in the reference's ontology taxonomy and also to the semantic mismatches, records defined at glossary building step. At the end of this round, as in the step before, participants are requested to present their opinion about the ontology defined, choosing one of the options: review (rev), reject (rej), accept (ok) or ignore (blank). If there is no agreement on the result this entire process is repeated again.

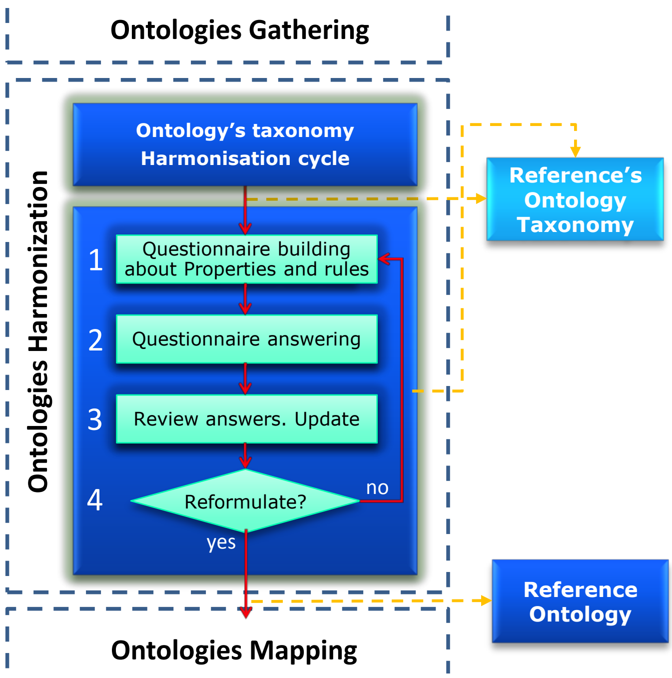


Figure 4.7 – Ontology's Contents Harmonization Cycle

4.3 *MENTOR Light Version*

MENTOR also provides guidelines to help on ontology building from scratch that due to its easiness comparing to the MENTOR fully objective, are considered its light version. It is accomplished by the regular first three steps plus one related to MENTOR phase 2: the ontology building step. Figure 4.8 illustrates their differences: Figure 4.8 a) shows MENTOR with its six steps; and Figure 4.8 b) presents the four steps of the MENTOR light version.

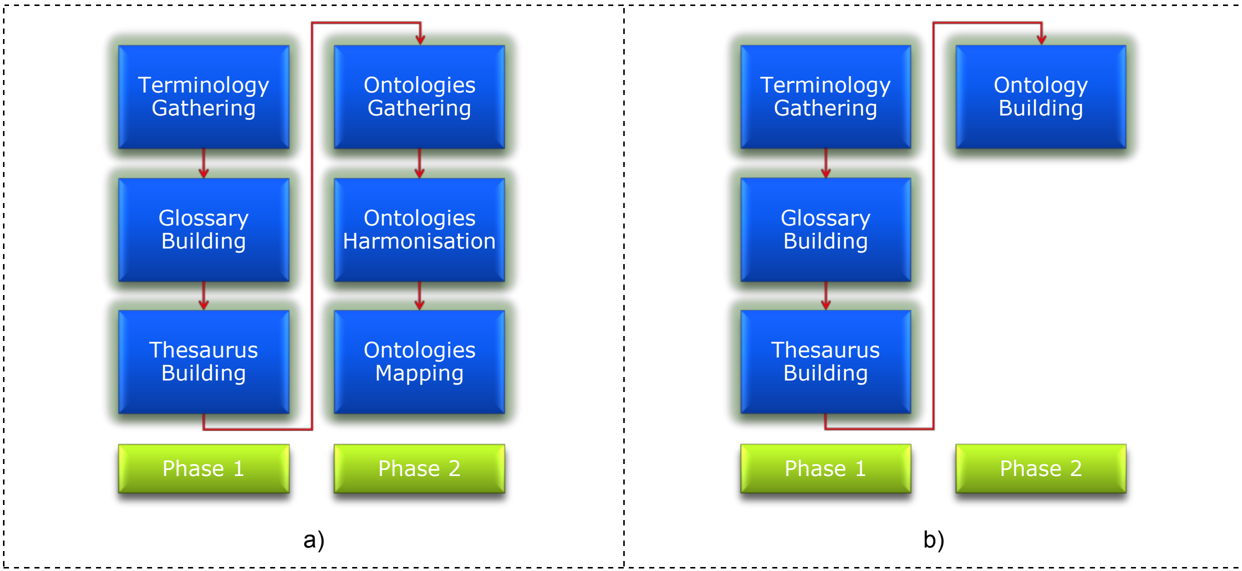


Figure 4.8 – a) MENTOR Full Version; b) MENTOR Light Version

The ontology building step addresses a normal ontology building procedure. Author proposes the use of BORO based method to build an ontology as a good guideline to help people to structure an ontology from a set of concepts (Figure 4.9). In this case, such concepts were defined in the first MENTOR phase specifically in glossary and thesaurus building.

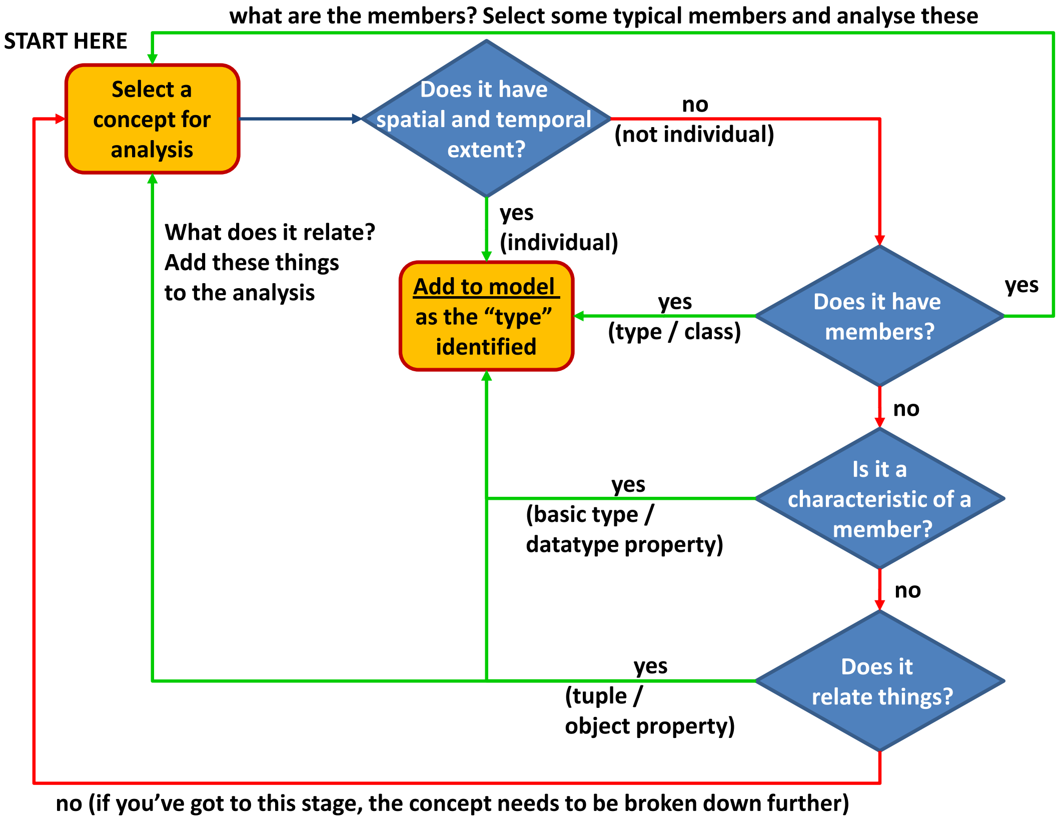


Figure 4.9 – BORO based method to build and ontology

The process starts by the selection of a concept (from the glossary) for analysis. If it represents a real "thing" with a spatial and temporal extent, it should be represented in the ontology as an "individual". If not there is a new question: "Does it have members?". If yes, the concept is a type or class, then the process should proceed to other concept analysis that is a member of the last concept analysed. If don't, should be analysed if the concept represents a characteristic of other members (concepts) and should be represented as a datatype property on the ontology. If the concept analysis fails in all the last three questions is because the chosen concept needs to be "broken down further". It should be clearly understood by the ontology builder people.

4.4 MENTOR's Approach to Mappings Establishment

Ontology mapping is the process that relates the vocabulary of two ontologies that share the same domain of discourse [286]. The process of defining mappings between ontologies is not an easy task and requires human support. The MENTOR uses the MO as the reference for mediating the mapping establishment and enabling subsequent 'mapping records' reasoning, for its interpretation. Having specific tools able to interact with MO, communities could build systems able to understand each other's representation format, without having to change their data and schema import or export processes.

4.4.1 **Mediator Ontology**

The MO has been built up as an extension to the Model Traceability Ontology defined by Sarraipa et al. in [287], and it is able to represent ontology semantic operations, like: the semantic mismatches found in the Glossary building step; the semantic transformations identified in the harmonisation process; the ontologies mapping and other ontologies operations (e.g. versioning).

Figure 4.10 depicts an abstract ontology life-cycle example with three ontology operations and three stages. Ontology N is the intermediate stage. From Ontology N is possible to make: a backward trace following an Ontology N entity to its related entity in the Ontology N-1 that is the root of this life-cycle, or a forward trace until the last version of the ontology (Ontology N+1). The Entity δ in the Ontology N is the result of a transformation that uses Entity η in the same ontology, which is related to the Entity α of the Ontology N-1. The Entity δ of Ontology N is in alignment to the Entity x of the Ontology N+1.

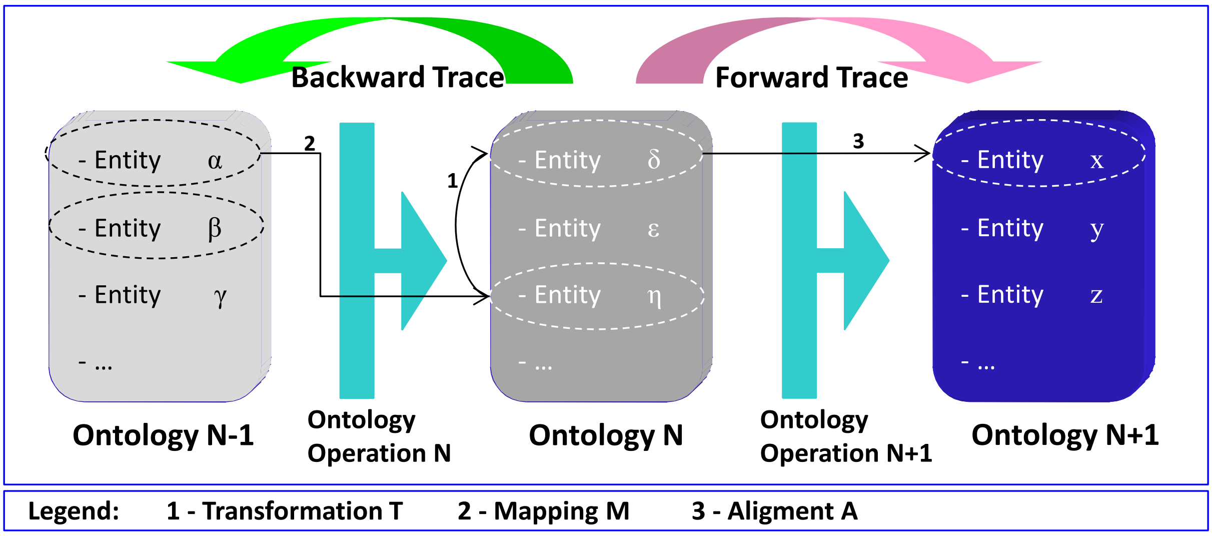


Figure 4.10 – Operations in a three stage ontology life-cycle

Thus, the MO is able to log ontology and entity operations in a way that is possible to trace changes in all the ontology life cycle. It addresses traceability as the ability to chronologically interrelate the uniquely identifiable objects in a way that can be processed by a human or a system. The mapping relations can be related to a traceability element, in such sense that a specific term defined in the reference ontology has a related one in the organisation member ontology, making possible a way to trace ontology elements. This way, the morphisms are modelled with traceability properties in a sense that they enable to store different versions of information model elements, as well as mappings between specific objects defined in a model or ontology A (*relating*) and objects defined in a model or ontology B (*related*). The structure of the evolved MO is presented in Figure 4.11 and described as follows.

The MO has two main classes: *Object* and *Morphism*. The *Object* represents any *InformationModel* (IM) which is the model or ontology itself and *ModelElements* (also belonging to the IM) that can either be classes, properties or instances. The *Morphism* basically represents the mapping tuple described in the section 2.3.2. It associates a pair of *Objects* (related and relating), and classifies their relationship with a *MorphismType*, *KnowledgeMappingType* (if the morphism is a mapping), and *MatchMismatch* class. The *Morphism* is also prepared to store transformation expressions in *Executable_Code* class that has been written using Atlas Transformation Language, which objective is to be used by several organizations to automatically transform and exchange data with their business partners as envisaged in [288].

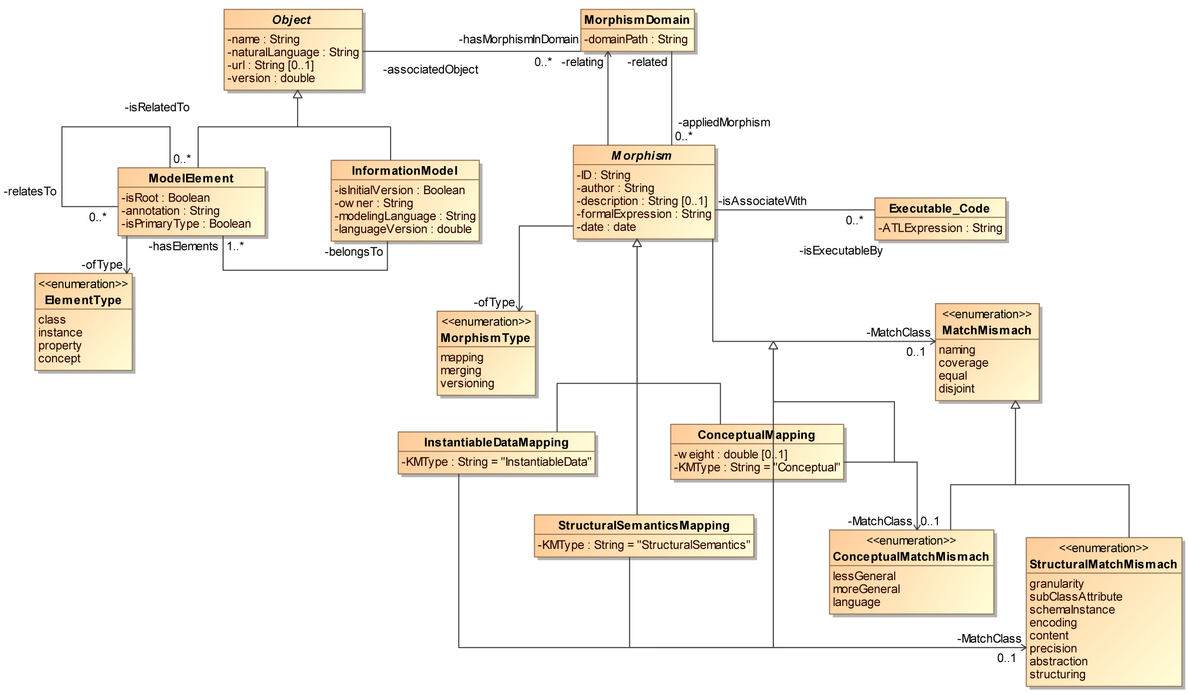


Figure 4.11 – MO Structure

Since the objective of the MENTOR methodology is to build an ontology that represents the global knowledge of an organisation domain, but keep working the old own enterprises ontologies, it is important to define how and in what processes all the ontologies involved should be maintained.

A typical situation refers to when new enterprises want to interoperate with an existent MENTOR compliant organisation. Here, the MO supports such integration, and furthermore both cases can be represented as a new stage in the reference ontology life. Reference ontologies must aggregate all the enterprise members' knowledge, when adding new enterprises to an existent organisation implies integration of all the new enterprise's knowledge in the organisation reference ontology. To achieve it, the new enterprise's knowledge has to follow the MENTOR methodology phases, where in this case there is already a glossary, thesaurus and a reference ontology established. Hence, the process is to follow the entire MENTOR steps taking into account the results from the first building reference ontology. Consequently, all the steps will have a lighter discussion or process in their outputs, where a slight refinement is expected, since the previous results have a bigger weight in the new reference ontology version consolidation.

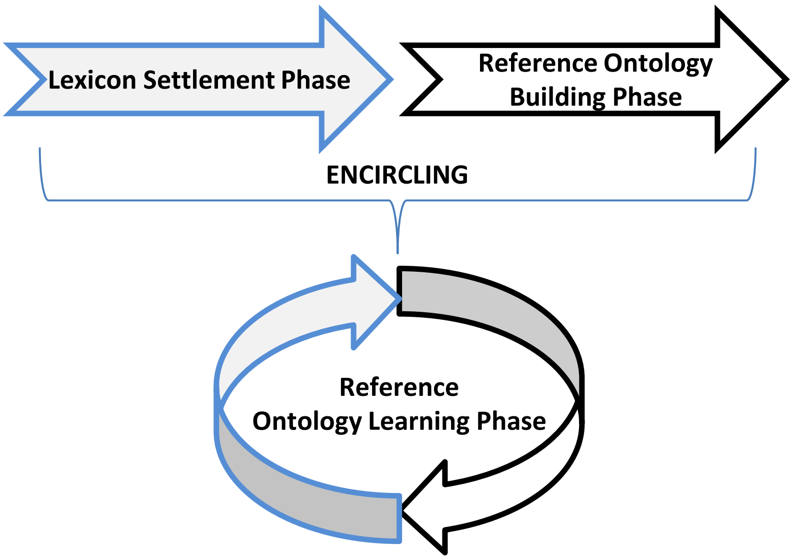


Figure 4.12 – MENTOR's third phase: OL

The mentioned maintenance procedures of MENTOR's ontology architecture are related to knowledge acquisition supported by a new OL phase. This phase implements the feedback mechanisms for a sustainable evolutional learning of the dynamic ontological system. Such a process is proposed to be the third phase of the MENTOR's methodology resulting in the encircling of the lexicon settlement and reference ontology building phases for a continuous OL (Figure 4.12). This mentioned knowledge maintenance is also addressed on the chapter 6.

4.5 MENTOR Demonstration

The simple choice of a "bolt" supplier by a mechanical engineer/designer, very often brings interoperability problems. Even after the appearance of standards in bolt specifications, suppliers keep defining various nomenclatures for their products and its associated knowledge. Thus, the need to align such product data sources is a priority to solve in the EI area.

4.5.1 **MENTOR Scenario**

The scenario presented in the following represents a real problem identified in the mechanical area or engineering. Two "bolt" suppliers, a retailer and a manufacturer compose it. Both of the enterprises have their own ontologies representing their own view of the domain. They intend to collaborate to build a common marketplace to sell their products. Such marketplace has to represent both own nomenclatures and at the same time to facilitate customers to find their products. As a consequence the marketplace should have a KB agreed by both stakeholders. Thus, the enterprises need to build a reference ontology to work as a mediator between their interactions and to their customers. To reach such objective, MENTOR methodology was followed (Figure 4.13).

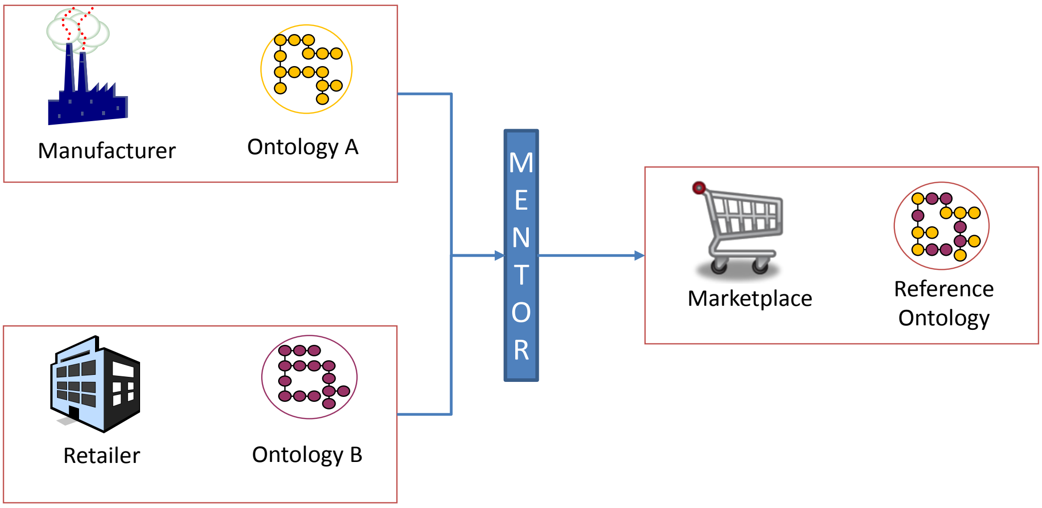


Figure 4.13 – MENTOR's use case scenario

4.5.2 **MENTOR Prototype Architecture and Components**

To put MENTOR into action, a prototype has been developed. Two different parts compose its architecture: the MENTOR server and the MENTOR Graphical User Interface (GUI) (Figure 4.14).

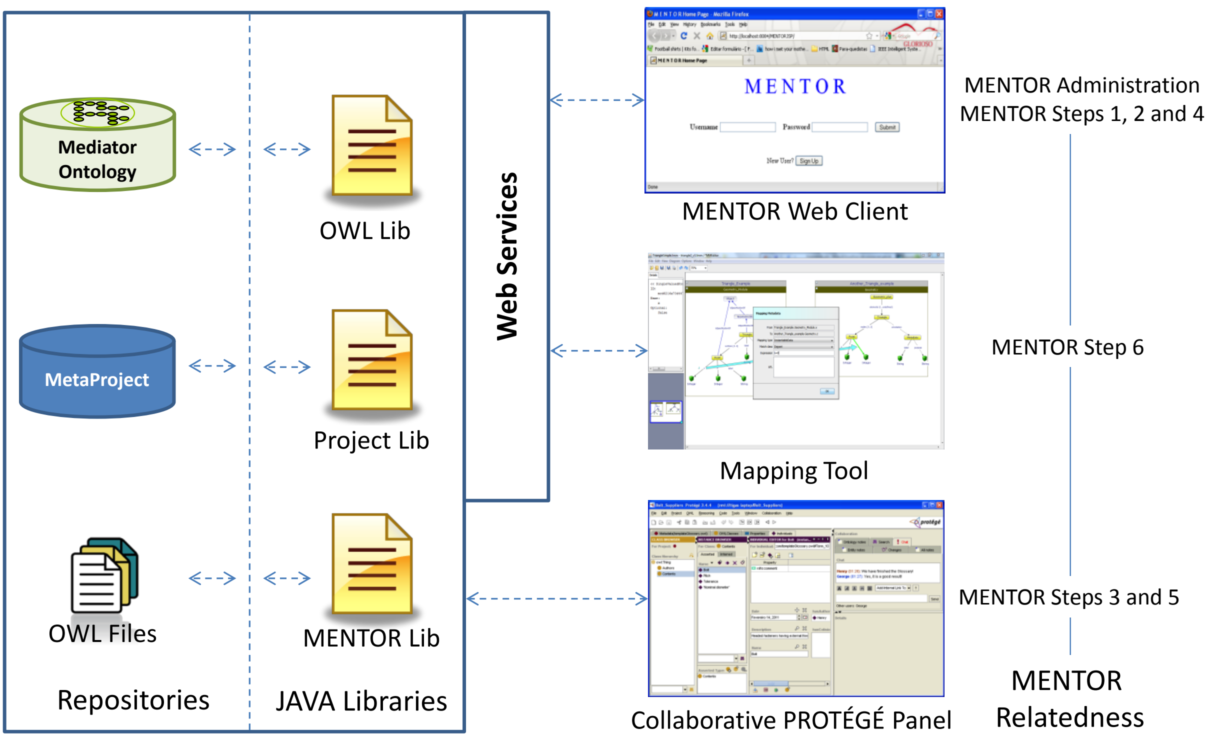


Figure 4.14 – MENTOR Prototype Architecture

Knowledge repositories and specific JAVA libraries able to provide interaction functionalities between the repositories and the GUI constitute MENTOR server. One of the used libraries is JENA - a Java API for OWL providing services for model representation, parsing, database persistence and querying ontologies. These libraries are used to implement a set of webservices, which are responsible for giving worldwide users the opportunity to use MENTOR services through a JSP based GUI. The knowledge repositories have the MO, the Metaproject file and a set of OWL files. The MO is used to record semantic mismatches and mappings. The Metaproject contains all the relevant data about users, groups, operations, permissions and projects, which establish the relation to the “OWL files” that represent the various files, saved in each MENTOR development process i.e. the glossary, thesaurus, and the reference ontology (Figure 4.15).

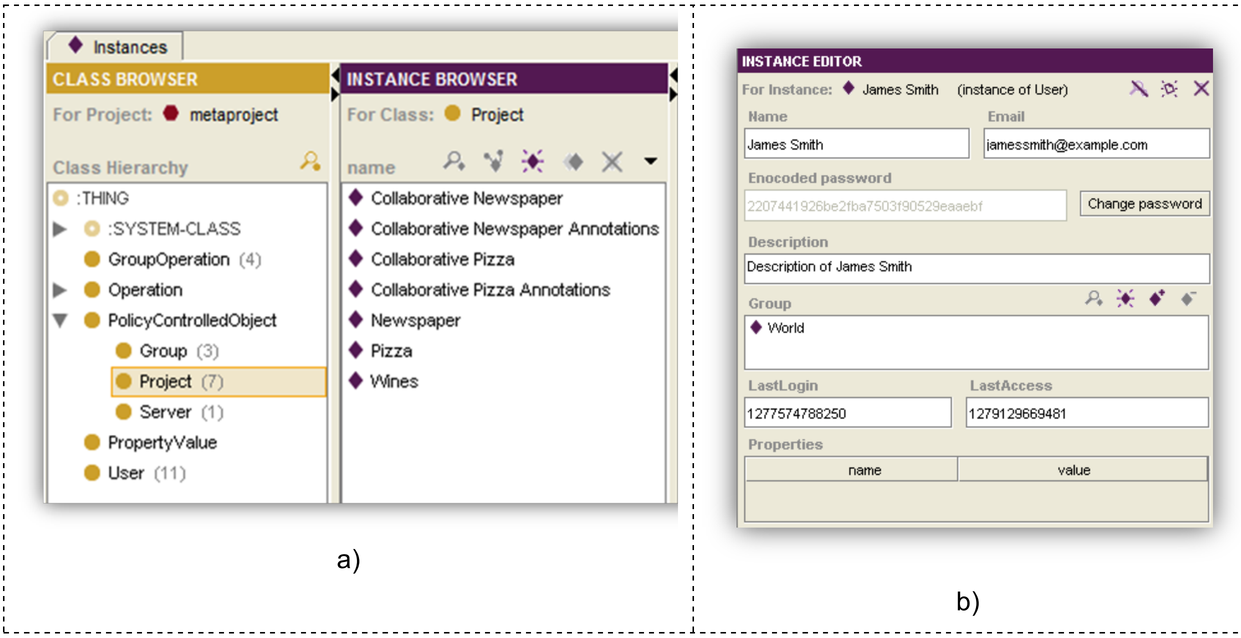


Figure 4.15 – Default a) Metaproject file; and b) user's instance

Since prototype developments are not fully integrated yet, MENTOR GUI is composed by three distinct components: the MENTOR web client; the Mapping Tool and the Protégé Collab.

MENTOR Web Client

MENTOR web client is the GUI used for the MENTOR administration and controls (Figure 4.16 b), providing also the necessary interface to all the users accomplishes MENTOR steps 1, 2 and 4. It provides a Login page (Figure 4.16 a)) to all users. Through it they can access to their private areas, where they have own information about the (ontology development) projects that they are enrolled to. There are two different profiles in MENTOR: the regular users and the administrators.

Administrators are able to create ontology development projects and manage their users and permissions. These permissions are recorded on the Metaproject through specific developed webservices. Administrators also control the various MENTOR steps, as for instance the final voting of the terms revision procedure (*Figure 4.3*).

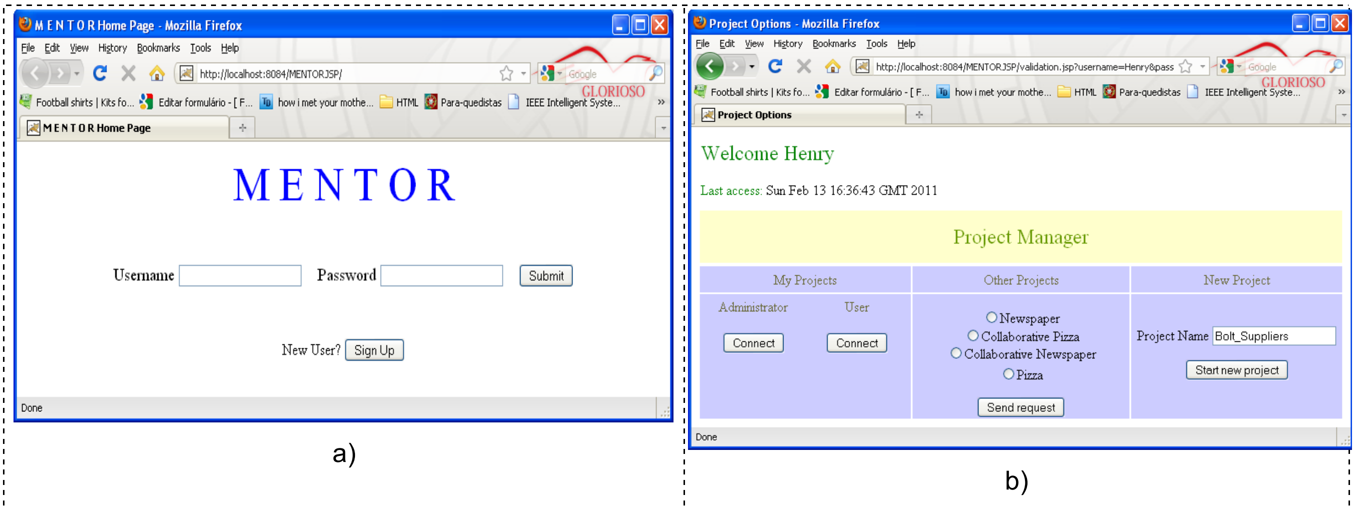


Figure 4.16 – MENTOR GUI a) Login Page; and b) Project Manager

Regular users are the main contributors to the ontology development procedure. They can enter to their areas to choose a specific ontology building project that they can participate. Thus appropriate interfaces will appear concerning the specific step task to which they have to contribute for. One task example is the terms revision task of the MENTOR Glossary Building step. Figure 4.17 illustrates the terms revision panel and Figure 4.19 a) presents the sequence diagram of such step interactions between users and the system.

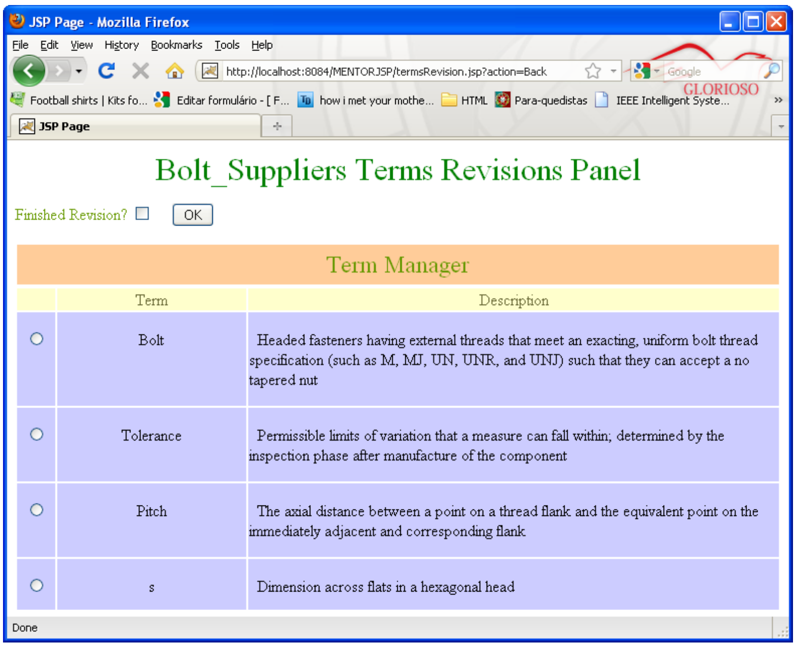


Figure 4.17 – Terms Revision panel

Mapping Tool

The mapping tool provides a graphical means to define the different kinds of mapping as described in 2.3.2, while storing them in the MO, thus it is used to accomplish the MENTOR step 6. As mapping definition is a complex and time-consuming task, the tool is capable of halting the process at any time without losing progress. JGraph has been elected and modified to read the input information model files and store morphisms at the MO. It is a widely used open source project for graph visualization and manipulation, similar to Microsoft Visio®, with good documentation and several examples. Features include a complete selection of layouts to automatically position the graph, many styles of shapes and edges, validation of connections, as well as an undo and redo manager.

Some adjustments were made to enable the interaction (mapping definition) between two different information models' graphs, and to become integrated the MO and to Language Independent Meta-Models (LIMM). JENA was used for the integration with the MO and a JAVA binder (JAXB) was included to allow the unmarshalling (interpretation) of LIMM files. LIMM serves as an abstract interface on top of enterprises' information models. Through its usage, becomes possible to abstract the technology and implementation details associated with the different modelling languages, and thus, enlarge the scope of users involved in a traditional mapping definition activity. In comparison to most modelling languages, it is intended to enable as little loss of expressiveness as possible, but at the same time, be simple and generic to support multiple language mappings.

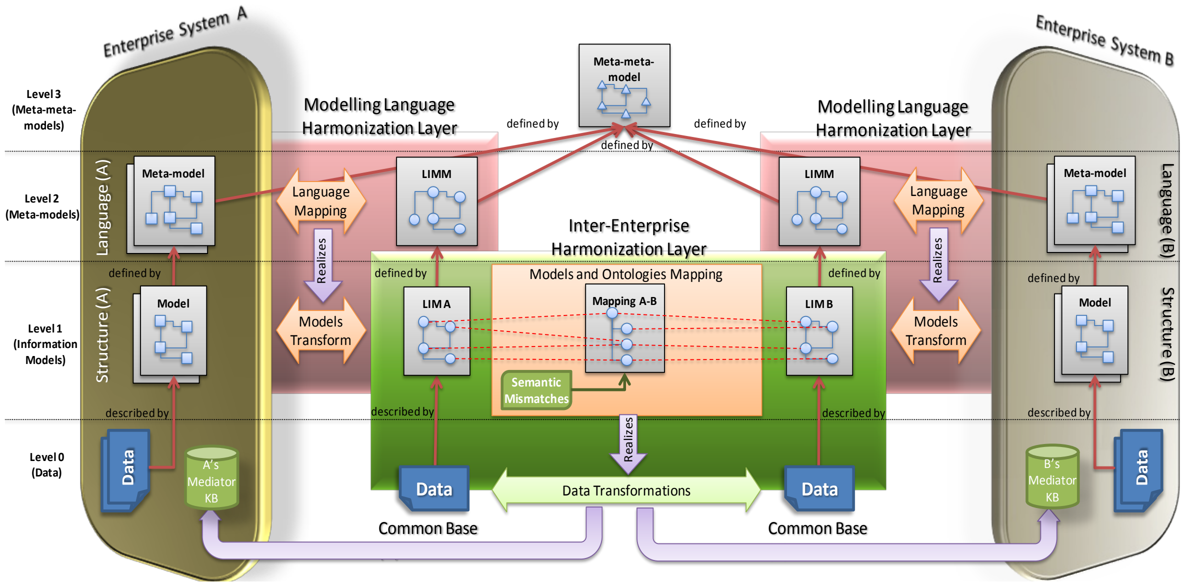


Figure 4.18 – MDA-based Framework for Interoperability Establishment [289]

LIMM has been developed in the GRIS-UNINOVA[31], and author has been participated mainly in its expressiveness testing specially in relation to ontology representation languages as OWL. LIMM is better described in Annex 10.1. LIMM is used in this tool to facilitate the mappings establishment between ontologies without obligate its organizations to use or translate such ontologies representation to a specific language (e.g. OWL). This tool implements the *MDA[32]-based Framework for Interoperability Establishment* defined by Agostinho et al. in [289], which solves interoperability problems associated to model languages transformations. Moreover, it also includes data transformations allowing the encapsulation of all the semantic discussions or agreements made through MENTOR. Thus, it enriches its Ontologies Mapping step. This *MDA-based Framework for Interoperability Establishment* implementation architecture (Figure 4.18) makes use of MDA's horizontal transformations to support the harmonization of modelling languages, models and data

[31] GRIS-UNINOVA – Group for Research in Interoperability of System at UNINOVA Research Center (www.uninova.pt)
[32] OMG's Model Drivel Architectures (MDA). www.omg.org/mda/

levels, within a platform independent context [289].

The left and right-hand sides of Figure 4.18 represent two different organization's information systems with different internal legacy models, where information is presented following the model- language-meta-model. The core of the architecture is focused on the middle part of the Figure 4.18, enabling two complementary layers, i.e. the modelling language harmonization layer and the inter-enterprise harmonization layer.

The modelling language harmonization layer (boundaries shared with the enterprises), is focused on the definition of mapping morphisms at the meta-model level, i.e. the modelling language used in each information model. It is therefore the layer realizing the transformation of models from one language to the other, which in our case, is used as an intermediate step for interoperability establishment. Enterprise system models, standards or even reference ontologies are transformed to their abstract interfaces (and vice-versa) using metadata descriptions (the LIMM) similar to the suggested in ISO/IEC 11179 Metadata Registries [290].

The inter-enterprise harmonization layer (centre), works sequentially after the first and is responsible for the model and semantics harmonization, defining mapping morphisms among the different abstract model interfaces. The process includes storing this knowledge in a MO (as presented at sub-chapter 2.3.2) replicated by the involved organizations, which serving as a standard during the mapping establishment will support the package for sustaining systems interoperability. The mapping tool GUI is mainly associated to this since it facilitates the interaction with the users for the mappings establishment. Note that "semantic mismatches" block presented on this area, represents the Glossary Building step output. It provides previous mismatches information related to semantic agreements established in previous MENTOR discussions, thus it could help users on the mappings establishment process.

Collaborative Protégé Panel

Collaborative Protégé as already described in the section 2.2.3, it is a Protégé plug-in able to perform track changes to ontology changes, which facilitates a distributed and collaborative manage of ontologies. It accomplishes other functionalities as chat, comments on ontology changes, voting, etc. The main feature of Collaborative Protégé is the ability to create notes attached to different things. This is the same idea as if someone would read an article and would add marginal notes on the paper. In the same way, the notes mechanism of Collaborative Protégé allows a user to create his own remarks (annotations) about a certain part of the ontology. This feature can also be used to discuss the ontology with other users either in stand-alone or multi-user mode [120]. This plug-in received some adaptations to have its comments and notes nomenclatures aligned to the various MENTOR task steps.

4.5.3 MENTOR's Application Demonstration

The demonstration presents the execution of the methodology showing how the information is processed on each step of the presented mechanical scenario.

During the whole execution of the methodology the admin has the control of the project. He decides when to move on to the next steps and keeps track of everything that is happening (as shown in the sequence diagrams - Figure 4.19). In this demonstration it is assumed that users Henry and George are the Retailer and the Manufacturer of Bolts of the scenario.

When all users are already connected to the project and ready to start contribution with their terms the project leader activates the first step. The generic representation of how the Terminology Gathering step is performed can be seen on Figure 4.19 a). The number of users or the amount of their contribution is not a relevant point for the system administration. Even with just one user it would be possible to build a domain ontology, the system permits from one to several users for doing so.

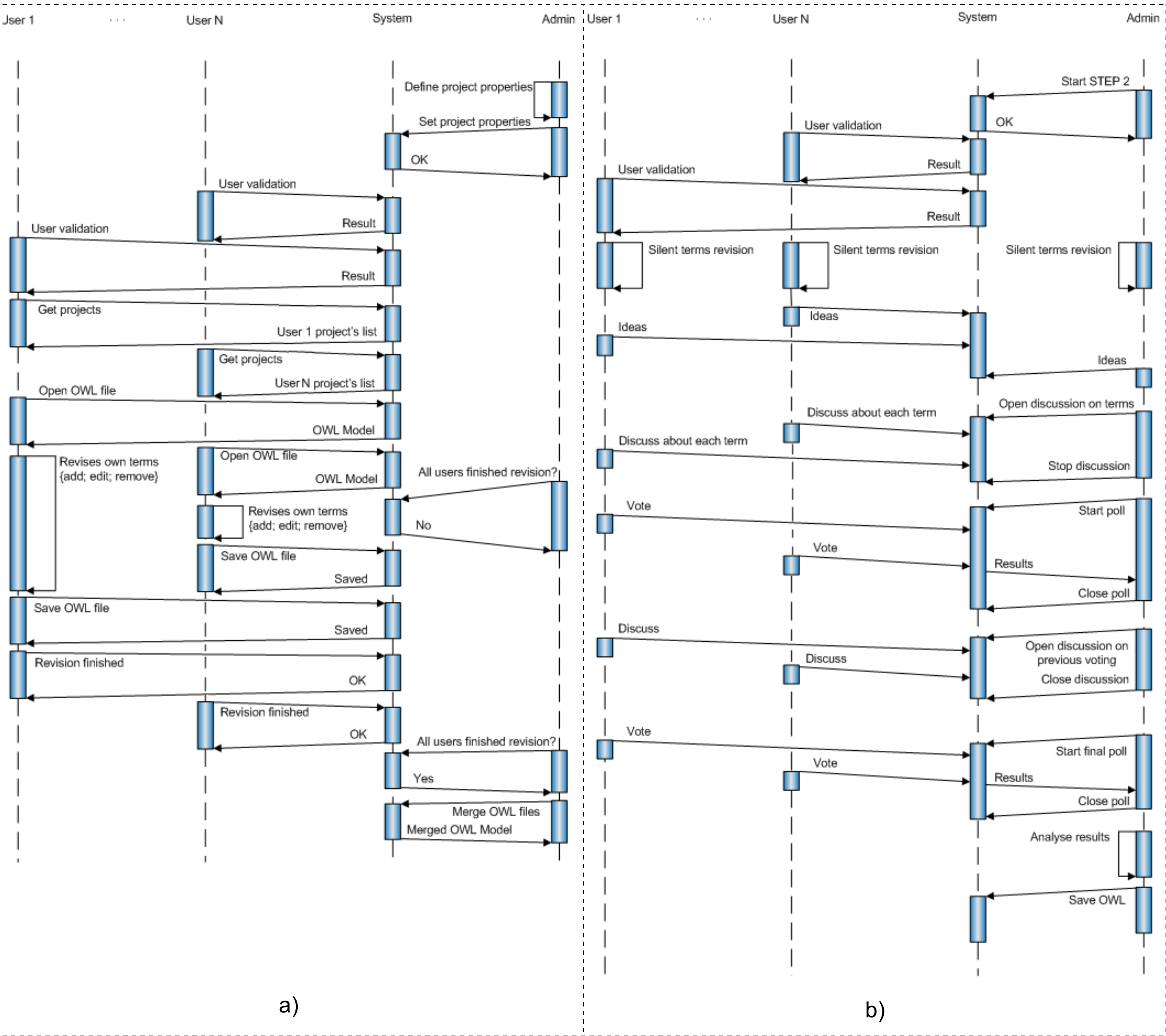


Figure 4.19 – Sequence Diagrams: a) Terms Revision Step; and b) Glossary Building Step

In this step there is no need of any specific QICM approach since there is no interaction between the project participants. All users revise their terms and definitions for themselves and when they finish they send a confirmation to the admin that proceeds to the next step when every member had finished the revision.

Each user introduce in the system the terms and definitions they use. Table 4.1 and Table 4.2 present an excerpt of the retailer and the manufacturer terms and definitions.

Ontology	Term	Definition
Retailer	Bolt	Headed fasteners having external threads that meet an exacting, uniform bolt threads specification (such as M, MJ, UN, UNR and UNJ) such that they can accept a no tapered nut.
	Nominal Diameter	The diameter of an imaginary cylindrical surface tangent to the crests of an external and (or) to the roots of an internal thread.
	Maximum Diameter	The maximum value acceptable for the diameter obtained from a predefined allowed upper deviation of the nominal diameter.
	Minimum Diameter	The minimum value acceptable for the diameter obtained from a predefined allowed lower deviation of the nominal diameter

Table 4.1 – Retailer Ontology Terms and Definitions

Ontology	Term	Definition
Manufacturer	Bolt	Term used for a threaded fastener, with a head, designed to be used in conjunction with a nut.
	Nominal Diameter	Diameter of an imaginary cylinder parallel with the crests of the thread; in other words it is the distance from crest to crest for an external thread, or root to root for an internal thread.
	Tolerance	Allowable deviation from a nominal or specified dimension, determining maximum and minimum material condition.

Table 4.2 – Manufacturer Ontology Terms and Definitions

When a user finishes revising its terms and corresponding definitions, an OWL file is stored in the server side for keeping the information for further steps.

Then when administrator confirms that everyone finished revising terms, he can set the status of the *Terminology Gathering* to *Finished*. It is now time to proceed to the next step, the *Glossary Building*. But before passing to the next step it is needed to merge all OWL files created on the previous step into a single OWL file, relating each term to its creator. In this single created OWL file are stated the collisions that exist between terms of different users. For example terms with the same name or terms where the definition is equal. On the presented case the terms *Bolt* and *Nominal diameter* from Henry collides with the terms *Bolt* and *Nominal diameter* from George, respectively. Figure 4.20 shows the OWL file created in the end of Terminology Gathering step and for each term it is saved the information of: *Name*, *Description*, *Creation Date*, *Author*, and *Collisions*. On the field *hasCollisions* it can be seen a collision between Henry's term Bolt and George's term Bolt.

Figure 4.20 – OWL file generated in the end of Terminology Gathering

The Glossary Building step starts after the administrator sets the status to *Started*. Then, the interactions between users starts supported by the system like presented in the Figure 4.19 b) sequence diagram. Such process has a round robin of ideas and a voting mechanism reaching in this case, the agreed glossary as presented in Table 4.3.

Ontology	Term	Definition
Reference	Bolt	Headed fasteners having external threads that meet an exacting, uniform bolt thread specification (e.g. M, MJ, UN, UNR, UNJ) such that they can accept a no tapered nut.
	Nominal Diameter	In a hexagonal bolt's head, is the dimension of the nominal diameter tangent to the flats (also expressed as the dimension across flats which correspond to the size of wrench to use). The diameter of an imaginary cylindrical surface tangent to the crests of an external and (or) to the roots of an internal thread.
	Upper Tolerance	Maximum value of allowable deviation from a nominal or specified dimension.
	Lower Tolerance	Minimum value of allowable deviation from a nominal or specified dimension.

Table 4.3 – Reference Ontology Terms and Definitions

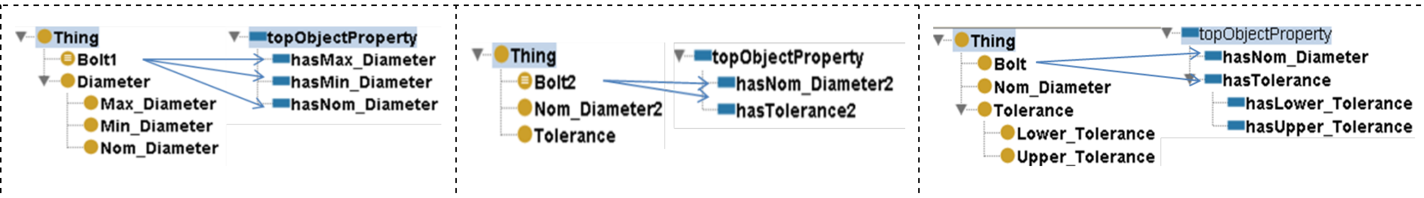
After having the glossary, it was defined the thesaurus. In such process it was identified the need of having the concept "Tolerance" from manufacturer to the thesaurus with the objective to be used as parent of the concepts "Upper Tolerance" and "Lower Tolerance". In this case, the thesaurus structure is equal to the reference ontology taxonomy (Figure 4.21 c)). The reference ontology accomplished with the insertion of instance properties, to "Bolt" class. This establishes a relation of this class to the "Nominal Diameter"; "Tolerance"; "Upper Tolerance" and "Lower Tolerance" concepts characteristics.

Note that during the harmonization phase it was identified that the "Maximum and Minimum Diameter" concepts from retailer ontology are obtained based on the following two equations that use the "Upper and Lower Tolerance" propreties.

upper tolerance + nominal diameter = maximum diameter

lower tolerance + nominal diameter = maximum diameter

Also it was identified the fact that the reference ontology distinguishes between "Upper and Lower Tolerances" while the manufacturer ontology does not. Upon obtaining the reference ontology the next step is to establish mapping tables between the proprietary ontologies and the reference ontology. The proprietary and the reference ontologies used to the mappings establishment of this scenario are represented in Figure 4.21 a); b) and c).



Retailer Ontology	Manufacturer Ontology	Reference Ontology
a)	b)	c)

Figure 4.21 - Used Ontologies

Through the mapping tool presented it was established the mappings between the retailer and reference, and between the manufacturer and reference ontologies respectively, as shown in Table 4.4 and Table 4.5.

ID		Retailer1_1	Retailer2_2	Retailer2.3_3	Retailer2.1_4	Retailer2.2_5
Melems = (a,b)	**a**	Retailer.Bolt1	Retailer.Diameter	Retailer.Nom_Diameter	Retailer.Max_Diameter	Retailer.Min_Diameter
	b	Reference.Bolt	Reference.Nom_Diameter	Reference.Nom_Diameter	Reference.Upper_Tolerance	Reference.Lower_Tolerance
KMType		Conceptual	Conceptual	Conceptual	Conceptual	Conceptual
Match Class		Equal	More General	Less General	More General	More General
Exp		$a = b$	$a \supseteq b$	$a \subseteq b$	$a \supseteq b$	$a \supseteq b$

Table 4.4 – Retailer Reference Mappings

ID		Manufacturer1_1	Manufacturer2_2	Manufacturer1.3_3	Manufacturer3_2	Manufacturer2.3_3
MElems = (a,b)	a	Manufacturer.Bolt2	Manufacturer.Nom_Diameter2	Manufacturer.Tolerance	Manufacturer.Tolerance	Manufacturer.Tolerance
	b	Reference.Bolt	Reference.Nom_Diameter	Reference.Tolerance	Reference.Upper_Tolerance	Reference.Lower_Tolerance
KMType		Conceptual	Conceptual	Conceptual	Conceptual	Conceptual
Match Class		Less General	Less General	Equal	More General	More General
Exp		$a \subseteq b$	$a \subseteq b$	$a = b$	$a \supseteq b$	$a \supseteq b$

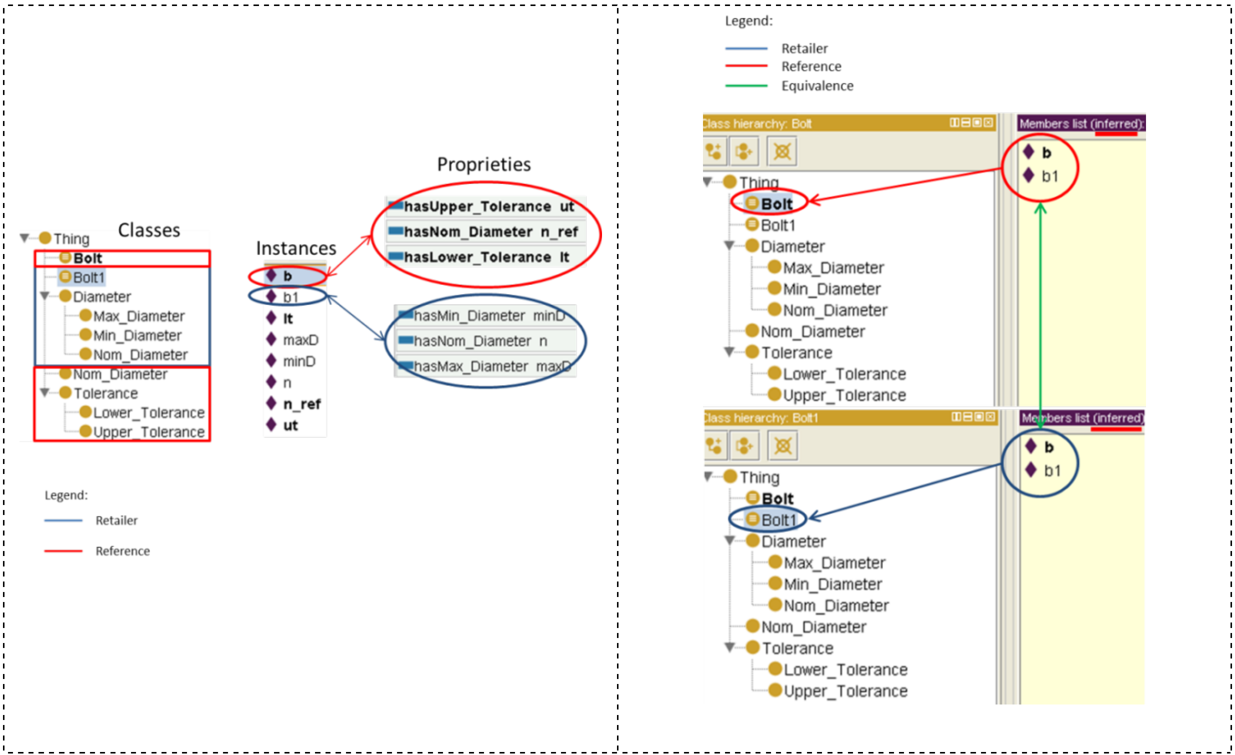
Table 4.5 – Manufacturer - Reference Mappings

However, having such mappings defined, it is very important to verify if the reference ontology indeed represents the knowledge gathered from the enterprises, and if any information model compliant with the reference ontology knowledge, is able to exchange data between the participants enterprises, without any loss of information independently to the direction that the data is transmitted to.

Thus, to accomplish such needs, author defined a framework for semantic checking of information systems, which is described in sub-chapter 8.1.1. The next examples feature the case of conceptual validation of multiple ontologies, namely between the retailer and reference ontologies and between the manufacturer and reference ontologies and vice-versa. The validation process followed a reasoning approach to check if the concepts are well represented in ontologies and aligned to all the participants' knowledge.

To accomplish that, the process starts by pairing the retailer and reference ontologies in the same KB. Then instances representing some of the concepts were created in the "Thing" class. These instances were created there to ensure that by reasoning the system puts them in their corresponding classes.

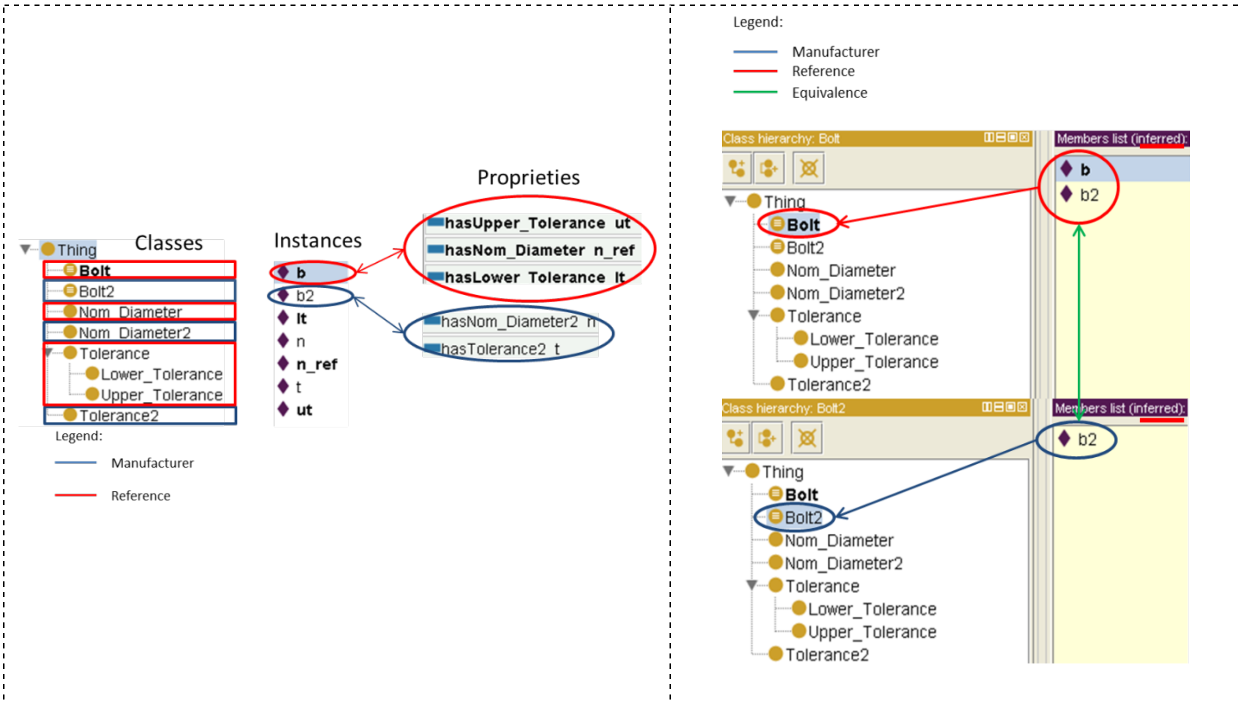
Consequently, as observed in Figure 4.22 a), two different types of "Bolt" instances (i.e. "b" and "b1") were created using properties of retailer and reference ontologies respectively. However, after running the reasoner, it was observed that both instances were indeed placed in the "Bolt" class of the retailer and reference ontologies (i.e. "Bolt" and "Bolt1") (Figure 4.22 b)). Therefore it can be concluded that the ontologies remained consistent and a bolt represented in the retailer ontology is semantic equivalent to a bolt represented in the reference ontology.



a) Before Reasoning b) After Reasoning

Figure 4.22 - Reasoning Example (Retailer and Reference Ontologies)

The next example is shown in Figure 4.23 denotes the manufacturer and reference ontologies. The principle of this example is the same as in the one before, meaning that two different types of “Bolt” instances were created within the “Thing” class (Figure 4.23 a)) and then the reasoning process was executed to verify if the instances were placed in their proper classes (Figure 4.23 b)).



a) Before Reasoning b) After Reasoning

Figure 4.23 - Reasoning Example (Manufacturer and Reference Ontologies)

Contrarily to the previous example, in this case, it is possible to observe some loss of information. Although both instances (“b” and “b2”) are represented within the reference ontology, the same cannot be said, regarding the manufacturer’s ontology, where only “b2” can be represented there. This is because of the way “Tolerance” definitions are represented by each of the ontologies. While the reference ontology distinguishes between “Upper and Lower Tolerances”, the manufacturers only define a single tolerance, assuming an equal value for “Upper” and “Lower”. This means that if different values for the “Upper and Lower Tolerances” are defined in the reference ontology then a conflict is created. Since the manufacturers’ ontology does not have such distinction, therefore will not know which value is the correct one, leading to possible inconsistencies in the KR. This loss of information is easily reflected in the mappings defined in the direction from the reference to the manufacturer, that are the same as the ones in Table 4.5 with the addition of the ones shown in Table 4.6.

ID		Reference3.1_1	Manufacturer3.2_2
MElems = (a,b)	**a**	Reference.Tolerance.Lower_Tolerance	Reference.Tolerance.Upper_Tolerance
	b	Manufacturer.Tolerance	Manufacturer.Tolerance
KMType		Conceptual	Conceptual
MatchClass		Abstraction	Abstraction
Exp		a = b	a = b

Table 4.6 – Reference - Manufacturer Conceptual Mappings

It is also worthy of remark that to aid in the reasoning process some rules were defined in SWRL. These rules serve the purpose of aiding the inference engine by providing it with additional facts and logical consequences that are based on the mappings defined earlier. Table 4.7 and Table 4.8 illustrate the rules defined in the first example and second examples, respectively and their purpose.

Rule	Purpose
Min_Diameter(?minD), Lower_Tolerance(?lt), Nom_Diameter(?n), Thing(?b), hasMin_Diameter(?b, ?minD), hasNom_Diameter(?b, ?n) -> hasLower_Tolerance(?b, ?lt)	If a bolt instance is defined as having a minimum diameter and a nominal diameter then it can be concluded that it also has a lower tolerance.
Max_Diameter(?maxD), Nom_Diameter(?n), Upper_Tolerance(?ut), Thing(?b), hasMax_Diameter(?b, ?maxD), hasNom_Diameter(?b, ?n) -> hasUpper_Tolerance(?b, ?ut)	If a bolt instance is defined as having a maximum diameter and a nominal diameter then it can be concluded that it also has an upper tolerance.
Max_Diameter(?maxD), Nom_Diameter(?n), Upper_Tolerance(?ut), Thing(?b),	If a bolt instance is defined as having a nominal diameter and an upper tolerance then it can be

hasNom_Diameter(?b, ?n), hasUpper_Tolerance(?b, ?ut) -> hasMax_Diameter(?b, ?maxD)	concluded that it also has a maximum diameter.
Min_Diameter(?minD), Lower_Tolerance(?lt), Nom_Diameter(?n), Thing(?b), hasLower_Tolerance(?b, ?lt), hasNom_Diameter(?b, ?n) -> hasMin_Diameter(?b, ?minD)	If a bolt instance is defined as having a nominal diameter and a lower tolerance then it can be concluded that it also has a minimum diameter.

Table 4.7 – SWRL rules defined in the retailer - reference example

The rules in the Table 4.7 explore the diameter and tolerance proprieties of the ontologies and proved to be invaluable to validate the semantic consistency of the ontologies. It is quite simple to conceive that bolts can have slight deviations regarding their diameters, so by defining a nominal diameter and upper and lower tolerances it is easy to conclude that the bolt has maximum and minimum diameters. The contrary is also true, if a nominal diameter for a bolt is defined as a certain value and the end product records a slight deviation either by excess or default then it is easy to conclude that the bolt has upper and lower tolerances.

Rule	Purpose
Tolerance2(?t2), Lower_Tolerance(?lt), Thing(?b), hasTolerance2(?b, ?t2) -> hasLower_Tolerance(?b, ?lt)	If a bolt instance is defined has having a tolerance then it can be concluded that it also has a lower tolerance.
Tolerance2(?t2), Upper_Tolerance(?ut), Thing(?b), hasTolerance2(?b, ?t2) -> hasUpper_Tolerance(?b, ?ut)	If a bolt instance is defined has having a tolerance then it can be concluded that it also has an upper tolerance.

Table 4.8 – SWRL rules defined in the manufacturer - reference example

These rules, in the Table 4.8 exploit the tolerance definitions of the manufacturer and reference ontologies. In this case it is assumed that if a “Bolt” is defined has having a “Tolerance” it can be concluded that it has both the same “Upper and Lower Tolerances”. However, unlike the previous conclusions, the contrary is not true, since the bolt can have different “Upper and Lower Tolerances” it is not possible to conclude that it has a single “Tolerance”. As a consequence this can lead to inconsistencies as it was explained before.

To sum up this scenario, Table 4.9 illustrates the information exchange between ontologies and whether this exchange resulted in a loss of information. There is a loss of information only in one case, from the reference to the manufacturer ontology. This means that the conceptual checking has failed in this case, since not all the knowledge represented in the reference ontology can be reproduced in the manufacturer ontology. As previously explained, this has to do with the tolerance definitions adopted by both entities. This loss was recorded from the reference to the manufacturer, what is to be expected when the mappings in this direction had a match class of Abstraction, which is a lossy semantic mismatch. On the other hand, no loss of information was recorded in the opposite direction, i.e. from the manufacturer to the reference. This is due to the fact that the tolerance concept of the manufacturer ontology is more general than the upper and lower tolerance concepts of the reference ontology, thus the information can be ‘split’ evenly between the reference concepts. For example if the tolerance is defined as being 0,5 centimetres then the reference assumes the same value for both the upper and lower tolerances.

Ontologies Information Exchange (From – To)	**Information Loss**
Retailer – Reference	No
Reference – Retailer	No
Manufacturer - Reference	No
Reference - Manufacturer	Yes

Table 4.9 – Identification of conceptual losses in information

Regarding the retailer and reference ontologies, no information losses were recorded in both directions since the concepts defined in each one are quite similar to one another.

MENTOR Results Application using a MDA-based architecture for transformation of models

The main idea of applying MENTOR to an organisation is to establish a semantic alignment between its actors or participants. Thus, as described in the scenario presentation, after have followed all the MENTOR methodology steps is time to apply its result to the exchange of information. In the scenario, both retailer and manufacturer used to represent their products and services in XML accordingly to the structure presented in their ontologies. Therefore, the MENTOR application testing is through a message exchange in a request for a product as depicted in Figure 4.25.

ID		Retailer2.1.1_1	Retailer2.1.1_2	Retailer2.2.2_3
MElems = (a,b)	**a**	Retailer.Bolt1.hasMax_Diameter	Retailer.Bolt1.hasNom_Diameter	Retailer.hasMin_Diameter
	b	Reference.Bolt.hasTolerance.hasUpper_Tolerance	Reference.Bolt.hasNom_Diameter	Reference.Bolt.hasTolerance.hasLower_Tolerance
KMType		Instantiable Data	Instantiable Data	Instantiable Data
MatchClass		Granularity	Equal	Granularity
Exp		“a = b + Nominal Diameter”	“a = b”	“a = Nominal Diameter- b”

Table 4.10 – Retailer - Reference Instantiable Data Mappings

There is still the need of establishing the mappings about the instance data of such information models. As in the previous examples, mappings were then also defined between concepts as being of the instantiable data types of such information models, in both directions, i.e., from the reference to the retailer and manufacturer and from the retailer and manufacturer to the reference. The mappings defined are illustrated in Table 4.10 and Table 4.11, respectively.

ID		Manufacturer3.2_1	Manufacturer3.2_2	Manufacturer2.1_3
MElems = (a,b)	a	Manufacturer.Bolt2.hasTolerance2	Manufacturer.Bolt2.hasTolerance2	Manufacturer.Bolt2.hasNom_Diameter2
	b	Reference.Bolt2.hasTolerance.hasUpper_Tolerance	Reference.Bolt2.hasTolerance.hasLower_Tolerance	Reference.Bolt2.hasNom_Diameter
KMType		Instantiable Data	Instantiable Data	Instantiable Data
Match Class		Granularity	Granularity	Equal
Exp		$a = b$	$a = b$	$a = b$

Table 4.11 – Manufacturer - Reference Instantiable Data Mappings

Both tables reflect the mappings in the direction from the retailer and manufacturer to the reference ontology. On the opposite direction mappings from the retailer to the reference remain in this specific case, the same as the ones in the reference to retailer (Table 4.10). On the other hand the mappings from the manufacturer to the reference are slightly different as shown in Table 4.12.

ID		Reference1.2.2_1	Reference1.2.1_2
MElems = (a,b)	a	Reference.Bolt2.hasTolerance.hasUpper_Tolerance	Reference.Bolt2.hasTolerance.hasLower_Tolerance
	b	Manufacturer.Bolt2.hasTolerance2	Manufacturer.Bolt2.hasTolerance2
KMType		Instantiable Data	Instantiable Data
MatchClass		Abstraction	Abstraction
Exp		$a = b$	$a = b$

Table 4.12 – Reference - Manufacturer Instantiable Data Mappings

After the successful mapping definition using the mapping tool presented before it is possible that the enterprise A LIMM instances can then be transformed to an enterprise B LIMM using the combination of the ATL rules specified for each mapping. The transformation is possible due to the mapping tool architecture, which implements the presented MDA-based Framework for Interoperability Establishment (Figure 4.18). It implements the transformation morphisms relying on a four level architecture that structures the relationships between meta-meta-models, meta-models, information models and data. The left-hand side of the Figure 4.18 represents the enterprise A model, using its own information modelling language, whereas the right-hand side represents the enterprise B internal model. Using a common meta-meta-model, such as the OMG MOF[33] it is possible to define the mappings among the meta-models at the level 2 of the MDA, which are the specifications of the modelling languages. With this, the transformation from any model to the desired format B at the Level 1 can be realized.

Given the context of MDA and MOF based meta-models transformation languages, the Atlas Transformation Language (ATL) is currently the largest user-base and has the most extensive available information such as reference guides, tutorials, programmers' forum, etc. It is the most used language to implement MDA based tools [291], having a specific Development Toolkit plug-in available in open source from the GMT Eclipse Modelling Project (EMP)[34]. Since the ATL can be applied to OMG meta-models [292][293], automatic model transformations at the information model level are attained if the mapping of level 2 (language level) is written in ATL [294].

Consequently, using the proposed architecture, the language mapping procedure is a manual process, but the language transformations are always automatic and repeatable. Considering that the number of languages used for information modelling is not so high, it is an acceptable cost since each mapping is done only once for each language, independently of the number of times it is executed.

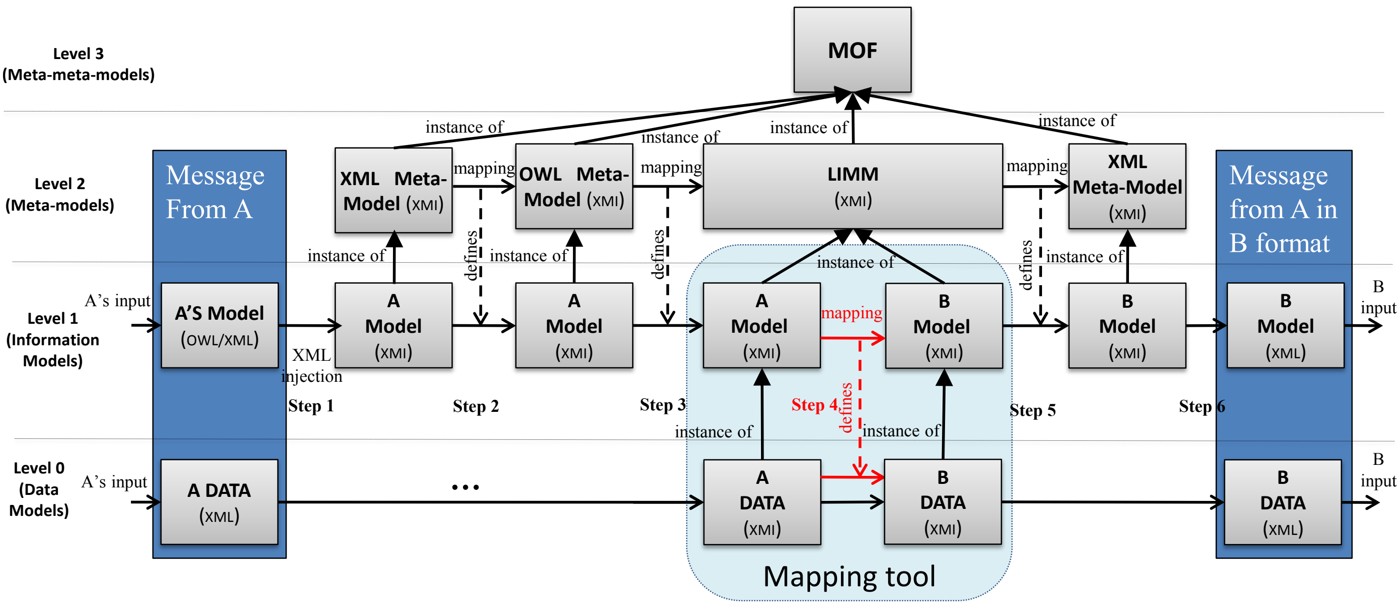


Figure 4.24 - ATL execution steps

The ATL transformation execution steps are represented in Figure 4.24. Its input models can be represented in OWL or XML, as in the case of the exchange message scenario presented in Figure 4.25. The steps represent the following actions:

[33] OMG Meta Object Facility (MOF). www.omg.org/mof/
[34] http://www.eclipse.org/modeling/

- **Step 1 (XML injection)** - The use of XML to be processed for the transformation tool needs to be represented in XMI. It can be natively injected by using a specific ATL (available in the used implementation framework);
- **Step 2 (OWL meta-model injection)** – the XML representation of the model introduced has to be injected to the OWL meta-model conforming model. This step use specific ATL defined a priori for transformations XML to OWL;
- **Step 3 (OWL to LIMM (LIMM model injection))** - OWL to LIMM injection uses the same procedure described in step 2. However now it uses ATL rules for OWL to LIMM transformations;
- **Step 4 (mapping through mapping tool)** – In this step, is when the mappings are made through the mapping tool. Such tool records the mappings accordingly to the mentioned tuple in MO, but it also askes user for specific ATL code to execute the mappings chosen. In this case the code is for LIMM to LIMM transformations. Note that actually the tool is not able to automatically produce the ATL accordingly to the mappings defined by the user. But it is desired that in the future the tool could make that automatically.
- **Step 5 (deserialise LIMM (XMI) to XML (XMI))** – the deserialise of LIMM to the B XML version, uses specific ATL defined a priori to such transformation type.
- **Step 6 (deserialise XML)** –This step is similar to the previous but this one uses a specific implementation framework solution able to transform models from XMI to XML serialised format.

The number of steps might cause the impression of great complexity. However, that is not the case and those have only been here included to guide readers in their implementations.

Finally, if you have a MO completely defined, which includes the ATL transformation code for each of its represented mappings, it can be used by a rule executor system. This system is then capable of executing model transformations accordingly to the mappings specified, becoming a Communication Mediator (CM). CM has a MO and in additional uses its information to mediate messages communications, executing model transformations when translations are needed. Figure 4.24 presents the core elements execution of a CM. In its centre, the mapping tool functionality complements CM, with the needed mapping information (including specific ATL execution code).

The scenario defined for MENTOR testing purposes was also able to test these complements. It shows how the information flows between a client, the CM and the enterprises (retailer and manufacturer). In this example, a client, through its system, orders a bolt product with particular specifications. As seen in the client message (Figure 4.25), the client specified a bolt with a nominal diameter of 10 and upper and lower tolerances of 0.2 and 0.1, respectively. A message containing these specifications is then sent from the client system to the CM in the reference ontology format. The CM then translates the message from the reference format, to both the retailer and manufacturers before relaying it to them. Converting from the reference to the retailer format is fairly straightforward. Based on the previously presented mappings, the CM only has to sum the “Nominal Diameter” and the “Upper Tolerance” to obtain the “Maximum Diameter”, subtract the “Lower Tolerance” to the “Nominal Diameter” to obtain the “Minimum Diameter” and the “Nominal Diameter” is the same for both.

However the case isn't so simple when translating from the reference to the manufacturer format. While the "Nominal Diameter" remains the same for both formats, the manufacturer, as explained in the previous section, doesn't distinguish between "Upper and Lower Tolerances", thus the CM has to assume one of its values, either "Upper or Lower Tolerance" (it's up to the system developer to choose which one or to reject such transformation with losses).

Based on the previous considerations the messages are translated to their target enterprises specific formats, having a specific structure (syntax) and semantics.

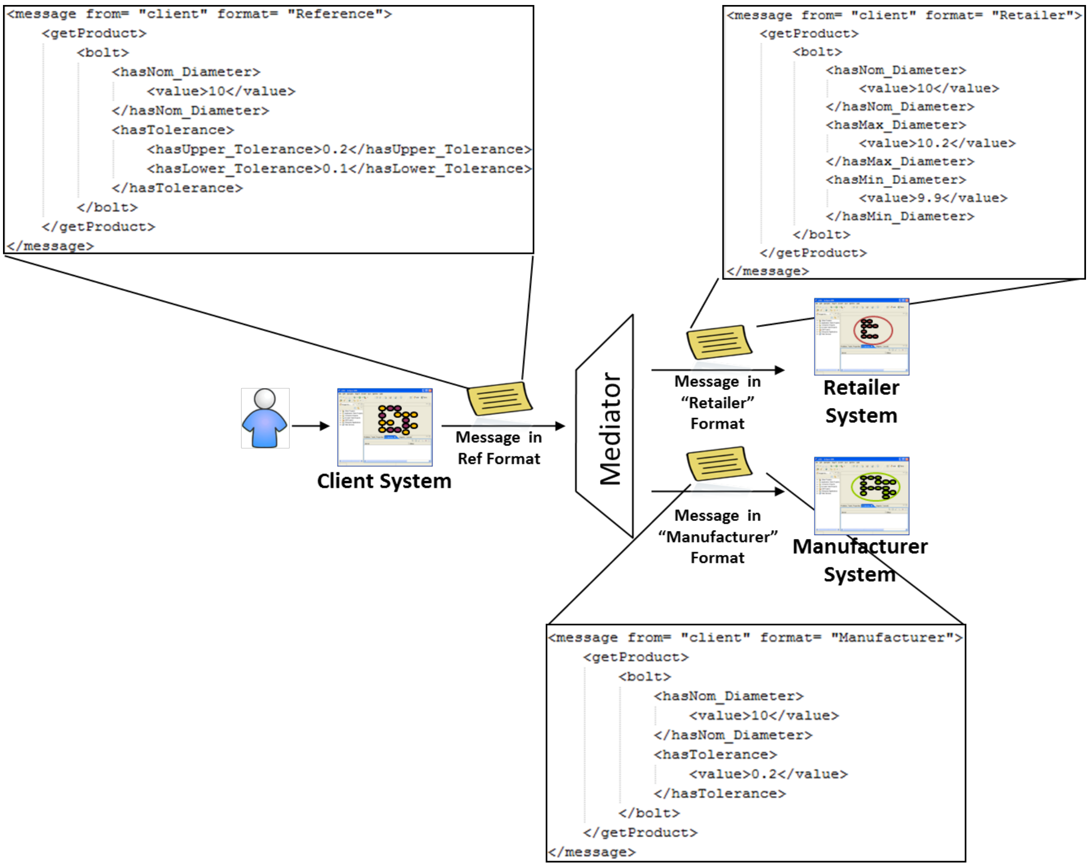


Figure 4.25 – Exchange Message Scenario

4.6 Conclusions

The proposed MENTOR methodology enhances inter and intra-organizational knowledge sharing, allowing its actors keeping their own ontologies or KR by producing a reference ontology in the domain. MENTOR brings together the building and reengineering of ontologies related to mapping competences.

MENTOR accomplishes traceability characteristics in its ontology maintenance approach mainly due to the specificity of the way that information of semantic changes is recorded in its MO. This information can be used to track any changes as going back to consistent previous ontology versions. As consequency of this capability, MENTOR enables dynamic and flexible seamless joining of enterprises to develop business in a network of partner organizations.

MENTOR was validated on this document through a case study related to the choice of a bolt in the mechanical area. However, in the past it was tested by the funStep initiative under the INNOVAFUN project (www.funstep.org) in their furniture reference ontology building, and presently has been tested under ALTER-NATIVA[35] project to support in an architecture development with the main goal to define curricular guidelines with technological support for higher education in the areas of language, mathematics and science, to support people in context of diversity.

These entire studies helped author to identify that tacit knowledge acquisition process isn't easy and clear. Sometimes discussions were too long, compromising the effectiveness of MENTOR. Thus, author developed and proposes for specific cases to use MENTOR as the collaborative ontology building methodology enriched with QICM to effectively improve the approach to elicit knowledge from business domain experts. Nevertheless, MENTOR is versatile; it is prepared for different situations as building an ontology from scratch; or working as a set of guidelines.

MENTOR was also prototyped as an element of an ISU to support on the setting up, establishing and facilitating the knowledge sharing among various enterprises and organizations [283].

[35] ALTER-NATIVA European Commission funded Research Project (DCI-ALA/19.09.01/10/21526/245-575/ALFA III(2010)88) – "*Referentes Curriculares con incorporación tecnológica par alas faculdades de educación en las áreas de lenguaje, matemáticas y ciencias, para atender poblaciones en contextos de diversidad*"

5 Knowledge Architecture for Multicultural Manufacturing Systems

The exchange of information and documents between partners often cannot be executed automatically or in electronic format as desirable, thus causing inefficiencies and cost increase [295]. This is primarily due to incompatibility problems among the several information representation structures used by the different software applications along supply chains and business networks [296].

With this diffuse range of systems, industry has had its development of trading partnerships restrained, e.g., inhibiting the shared fabrication of products, software solutions. These barriers are real factors that prevent innovation and development. Therefore, standardization rapidly became an evident priority, and several dedicated reference models covering many industrial areas and related application activities, from design phase to production and commercialization, have been developed enabling industrial sectors to exchange information based on common models [297].

Together with standards development, interoperability solutions have enabled a smooth progress of manufacturing systems to a next phase, where flexibility, intelligence and reconfiguration should be reached. The 'intelligence' concept becomes more relevant because of the need to maintain effective and efficient manufacturing operations with minimum downtime under conditions of uncertainty [298]. Intelligent is taken to mean advanced and efficient manufacturing technologies, management and procedures [299]. Therefore, one way to reach such Intelligence is exploring the use of formal ontologies as a way of specifying content-specific agreements for the sharing and reuse of knowledge among software entities [30].

Today, the simple use of standards as AP236 (Application Protocol 236)[300], commonly known as the Standard for the Exchange of Product model data (STEP) for *Furniture catalog and interior design*, or any other STEP Application Protocol, does not solve all the EI problems. Indeed, typically each stakeholder has his own nomenclature and associated meaning for his business products. Therefore, all business partners still may not understand the information exchanged, in spite of sharing the same structure.

Having great influence in the supply chain, namely in sectors characterized by large enterprises (which is not the case of furniture), manufacturers could force their own terminology. However that would cause chaos for their partners as they would have to deal with as many terminologies as the manufacturers they work with. Besides, this would reduce the advantages of using standards [296]. Also, more and more the global market is costumer driven, which means that retailers and marketplaces tend to adopt a terminology that costumers are familiar with, and in turn use it down the supply chain to satisfy customization requests.

Therefore, also the manufacturing organizations must be worried with their integration with the rest of the supply chain, and the manufacturing processes are dependent on the costumer procurement.

This interoperability need is related to the semantics of the contents exchanged, and the author, under his participation to the funStep initiative, proposed the semantic enrichment of the furniture product data as a solution. The main objective is to organize the knowledge associated to the furniture products in order to enable a full understandable business messages and catalogue exchange.

This chapter takes the furniture industry example to propose a knowledge architecture for semantic enrichment of standard-based electronic catalogue data. It is built upon different KREs, namely a dictionary, a thesaurus, a reference ontology in the furniture domain and the AP236 standard itself. Together, these KREs establish the syntax and the lexicon to be used in the furniture domain. Each concept has its own definition translated to different languages, and some of the concepts are classified as the reference ones to be used by the community in their information exchange. Specific links between all the KREs enable the establishment of ontology mapping solutions, so that enterprises can keep internally their own terminologies and classification systems, and still be interoperable with their business partners.

5.1 Funstep Knowledge Organisation

Data can exist in multiple ways, independently of being usable or not. In the raw format, it does not have meaning in and of itself. However, information is data that has been given meaning by way of relational connection to a context [301]. Still, in information, this "meaning" can be useful for some, but not necessarily to all. As Bellinger et al. in [48] define, information embodies the understanding of a relationship of some sort, possibly cause and effect. Thus, people might "memorize" information (as less-aspiring test-bound students often do), but still be unable to understand it because it requires a true cognitive and analytical ability, i.e. knowledge. In other side, as already described, Nonaka et al. in [276] defines two kinds of knowledge: 1) Tacit knowledge; and 2) Explicit knowledge. It is around these knowledge types that knowledge is handled by the funStep knowledge organisation.

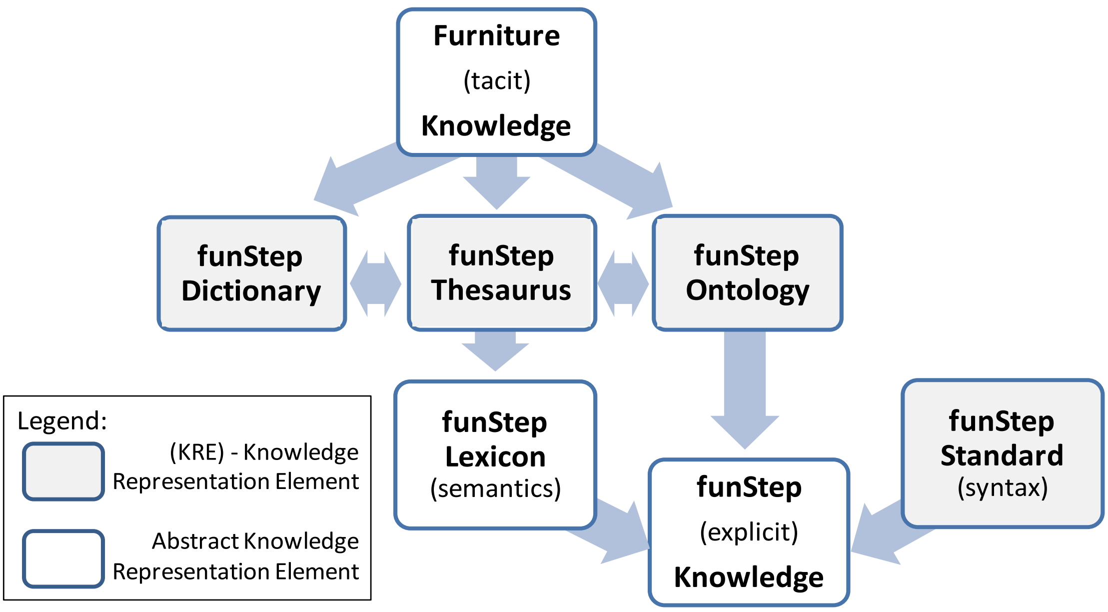


Figure 5.1 - funStep knowledge organisation

funStep Interest Group is a group focused in helping furniture business stakeholders in solving software issues related to its business information exchange and promoting its solutions. Thus, funStep endeavours to gather the tacit knowledge that furniture domain stakeholders hold into machine interpretable KBs. Closing this gap, funStep will obtain the (explicit) knowledge which should be stored in a structured organized way, where syntax and lexical semantics are integrated [302]. For reaching that objective, author proposed to integrate the funStep standard (AP236) with the reference funStep Lexicon, which embodies the reference concepts and semantics, and with a funStep ontology, which embraces product classification to its related properties. This leads to the funStep knowledge organisation (Figure 5.1). Therefore, the integrated knowledge is composed by four KREs (Figure 5.2): the funStep Ontology; the funStep Thesaurus; the funStep Dictionary, and the funStep AP236 ISO Standard.

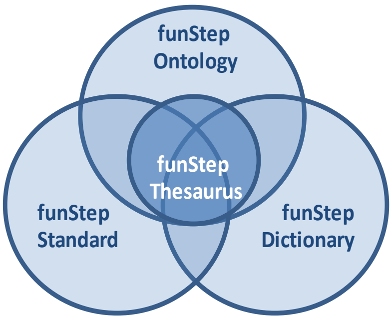


Figure 5.2 - funStep KREs domains intersection

For a good explicit KR, it is needed to have significant input from the tacit source (i.e., domain experts). Thus, such characteristic requires a knowledge architecture enabling the management of the evolution between the KREs. The evolution of the first three KREs leads to the funStep Lexicon establishment, which is an abstract KRE in the sense that it is composed by thesaurus contents, i.e. concepts and definitions. On the other hand, the funStep explicit knowledge KRE is another abstract KRE since it is composed by the addition of the funStep Lexicon with the ontology and the standard itself. In conclusion the funStep explicit knowledge represents all the furniture machine interpretable knowledge where the funStep dictionary and the thesaurus are supporting KREs to the funStep Lexicon establishment and maintenance.

Each one of these KREs has a particular scope, role and different objectives in the overall funStep knowledge architecture, where their focus domains intersect each other partially. The funStep thesaurus domain is totally included in the funStep ontology, while the other KREs have particular information that is not totally shared.

5.1.1 **The IPB Use Case**

There is main five stakeholders in the furniture supply chain, which are: suppliers, manufacturers, retailers, e-marketplaces and interior designer/architects, whose characteristics and relationships lead to different communications implementation requirements. According to Sendall et al. in [303] the analysis of a use case is one way to verify the system's functional requirements. Thus, it was considered an International Product Business (IPB) use case, aiming to represent situations, which could facilitate the description of how the knowledge is handled in an international plan of the furniture business.

In IPB, the globalisation phenomena and the international product business situation drove the author to the "leather couch" scenario analysis. This scenario is related to a "leather couch" request for quotation where a customer chooses, in a furniture catalogue, a foreign product (Figure 5.3).

The customer starts by asking the retailer for a "leather couch". The retailer shows the customer a set of catalogues with various types of "sofas". The customer looks at the catalogue and chooses the model "XPTO". However, in the catalogue it is not described if such model is available in "leather". Thus, the retailer has to ask manufacturer of the chosen model if it can be made in "leather". Since the model's manufacturer is in French, the request has to be translated. The retailer sends a fax with a request for quotation of a "XPTO en cuir" (translation of "in leather" to French). The retailer asks the customer for his contacts (e.g. email; phone; etc.) in order to be in touch.

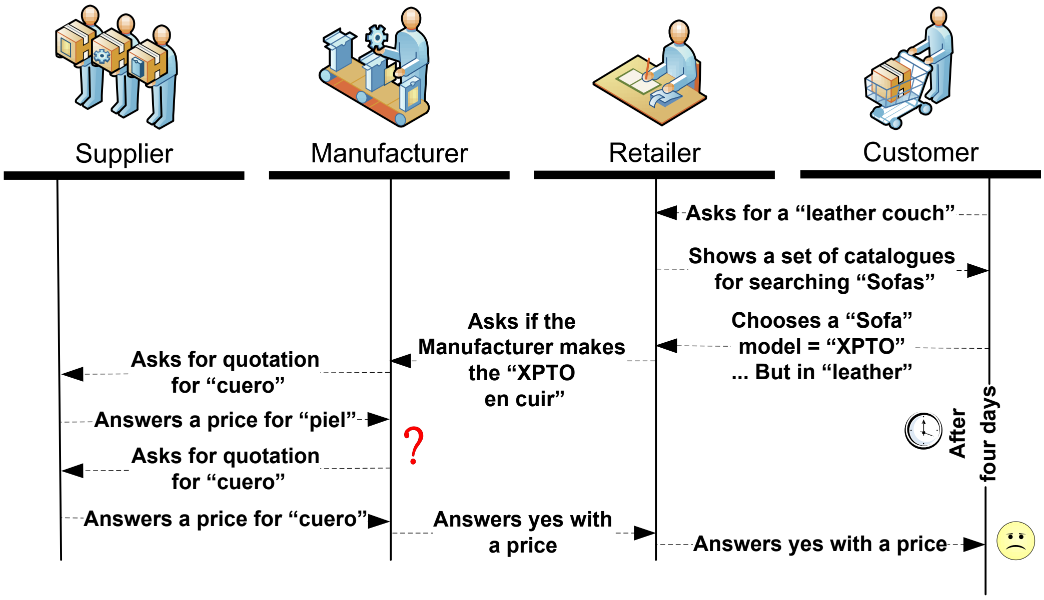


Figure 5.3 – Furniture case study scenario

The manufacturer does not have "leather" in stock, consequently he has to contact his supplier. Since his "leather" supplier is Spanish, he translated the request to "cuero" (translation of "leather" to Spanish). The supplier answered his request. Nevertheless, the manufacturer did not believe in the answer because the product description said "piel" (Spanish synonym word for "cuero") instead of "cuero". When the Spanish supplier received the reply asking for confirmation, he laughs because of the supposed misunderstanding that the French thought was happening. In order to avoid further interaction problems, the Spanish answered again using the "right" word – "cuero". Once the manufacturer received the quotation, he replied with the quotation for producing a "leather XPTO" sofa. After four days the customer received the answer.

The funStep initiative under the SMART-fm[36] project already defined a standard, namely AP236, for the exchange of data that could be used in this use case scenario communications [300]. Nevertheless despite the information exchange with AP236, semantic interoperability has to be

[36] SMART-fm. EC project Nº. IST-2001-52224. A standards-compliant framework to support complete integrated product life-cycle information management and electronic commerce for the furniture manufacturing (fm) industry, in the advent of the smart enterprises.

improved. From a first analysis to the presented use case scenario, it was emphasized the following set of statements that describe the necessity of having a structured furniture knowledge organization:

1) The retailer knows that “couch” is equivalent to “sofa”, and the catalogues containing these products mostly designate them as “sofas”. Thus, the reference word is “sofa” instead of “couch”.

2) The retailer only shows to the customer the catalogues that have sofa models. This implies a catalogue selection based in a classification item – “sofas”.

3) Once the customer wants the model “XPTO” in “leather”, the retailer had to translate this request to the language of the chosen model’s manufacturer.

4) In the interaction between the manufacturer and the supplier, it was identified the need of having defined reference terms and concepts for each translated language.

5) Since the retailer is English and the manufacturer is French, the quotation was converted to the right currency.

For each of the statements described it was defined a set of requirements that was taken in consideration in the process of reaching the desirable semantic interoperability in the furniture business interactions. Such requirements were:

1) To have a domain dictionary for getting help on choosing synonyms;

2) Establishment of a set of reference terms in the domain to be used in the business communications;

3) Define an ontology for products classification in order to enable knowledge reasoning;

4) To have a multilinguism dictionary to be used in translations;

5) Define for each language a set of reference domain terms for business interactions;

6) A way to enable data transformations, as for instance in currency conversions;

As conclusion of this scenario analysis, it was defined the need of having a multicultural knowledge-based architecture aligned to the funStep knowledge organization and able to match these requirements in the AP236 semantic enrichment process, which will facilitate the full information interoperability establishment in the furniture business. In addition, it was verified that from a business point of view, enterprise’s competitiveness is to a large extent determined by its ability to seamlessly interoperate in a collaboratively way with others.

5.2 The Funstep Multicultural Knowledge Architecture

Semantics is the study of language units meaning and their combinations [304]. Therefore, semantic enrichment is the act or process of adding specific meaning elements to some KR structure in a domain, to help on the information clarification [305]. The funStep standard semantic enrichment is characterized as being performed at two basic levels [306] [287]: 1) Terminological annotation, by using the terms identified in the Lexicon (Thesaurus); 2) Semantic annotation, by using concepts and expressions drawn from the Ontology.

Both levels are conducted through a multicultural knowledge architecture based on the KREs identified: the funStep dictionary, thesaurus, ontology and standard data. Matching the requirements identified in the use case presented, the author's proposal for the multicultural funStep architecture (Figure 5.4) is built up on web services technology enabling interoperable open services over the Internet between the funStep knowledge client systems and the funStep knowledge server [1]. Three parts compose this architecture: 1) natural language dictionary server; 2) funStep knowledge server and 3) funStep knowledge clients.

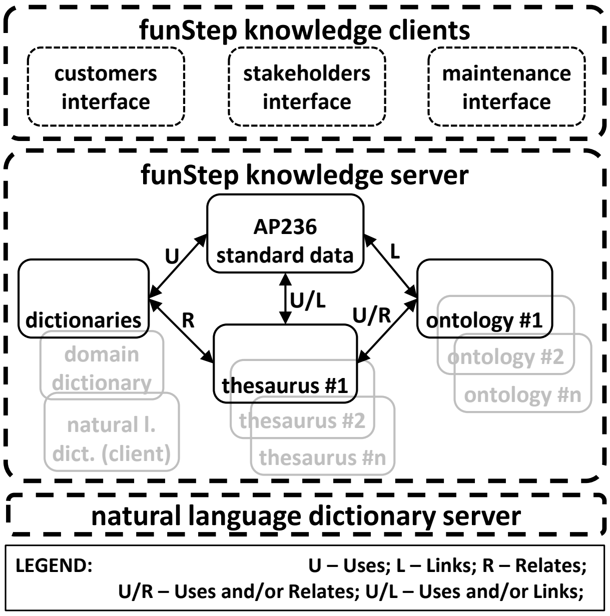


Figure 5.4 – funStep Multicultural Knowledge Architecture

5.2.1 The funStep knowledge server

The "funStep knowledge server" is composed by the four funStep KREs described in section 3, whose relationships enable the semantic enrichment of the standards data.

The AP236 standard data has links to the ontology for products classification, and to thesaurus concepts for external annotation to its data representation elements. This is made as explained before for the External Classification example (presented at sub-chapter 5.3). Thesaurus concepts and dictionaries terms (from domain and natural language dictionaries) are used in the knowledge server data representation and multilinguism translations. Some of the terms and meanings in the domain dictionary are related to thesaurus elements, which facilitate the retrieval of the related reference concepts, e.g. the "couch" concept was replaced by the reference concept "sofa".

The ontology is closely related to the thesaurus, sharing a very similar structure. This enables the usage of concepts definitions from both KREs for semantic clarification and to establish a relationship between them. Therefore, the funStep Ontology could be easily translated taking as source the thesaurus for the target language.

The funStep Dictionary

The funStep dictionary supports a multilingual collection of terms, thus enabling an enriched coordination between international partners. Also, the terms are associated to a description, synonyms, other related terms, and optionally multimedia items, such as images, sounds or videos. Continuing with the "leather couch" scenario, Figure 2.5 represents a view of the dictionary with the meaning of the term "Sofa"; its picture; its related words – in this case it is "couch"; and some translations (e.g. "canapé" in French; "Sofá" in Spanish and Portuguese).

The funStep Ontology

The funStep ontology main objective is to represent all the knowledge associated to furniture products, enabling semi-automatic classification. The funStep ontology started to be developed by the funStep initiative during the COFURN[37] European Project, based on search criteria and codes which were used than for furniture products classification in electronic commerce. Afterwards, during the SMART-fm project, it was upgraded through a harmonization process from existing taxonomies of different fields such as e-marketplaces and e-commerce websites [286].

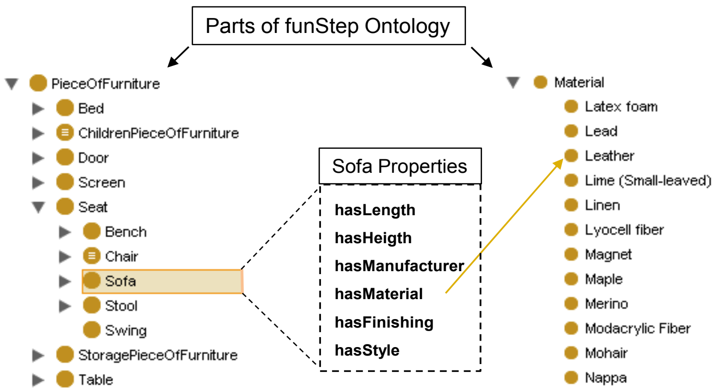


Figure 5.5 – Subset of the funStep Ontology taxonomy

Figure 5.5 depicts a subset of the funStep ontology taxonomy emphasizing the furniture product characteristics mentioned in the "leather couch" scenario – classification as "leather sofa".

The funStep Thesaurus

The basic lexicon establishment is reached by the development of a thesaurus on the domain – the funStep Thesaurus. This is composed by a set of domain reference terms and concepts, clustered on the basis of their similarity, organized by means of semantic relationships (e.g., equivalence, subsumption, generalization, disjunction), thus enabling a better retrieval process of semantically related terms [101]. A thesaurus can serve as a controlled vocabulary where terms are constrained to

[37] COFURN. EC project Nº. IST-2000-25183. CO-operation for consensus, standardisation and interoperability to support e-com services in the FURNiture sector.

its domain-specific meanings, avoiding the problem of ambiguity [307]. The funStep thesaurus envisages a multi-national scope of vocabulary, where terms with the same meaning coexist in multiple languages. Multilingual thesauri can be used to translate queries, by expanding the query to one or more target languages [308]. Still in the "leather couch" use case scenario, the multilingual thesaurus addresses the definition of the related words "Leather", "Cuir", and "Cuero".
The construction of a multilingual thesaurus typically begins with the analysis of business messages, costumer search requests, common domain documents and other knowledge sources [309]. Therefore, the thesaurus can be seen as a collection of parts of the dictionary, ontology and the standard itself (Figure 5.2).

Again, following the "leather couch" scenario, a subset of the multilingual thesaurus, can be as follows, where the last three concepts are linked each other since they represent the same meaning exactly in English, French and Spanish (Figure 5.6):

- **Sofa**: A sofa is a long upholstered seat typically with a back and arms for the comfortable seating of more than one person.
- **Leather**: A material made from animal skins treated by a special process.
- **Cuir**: Matériaux fait à partir de peaux animales traitées par un processus spécial.
- **Cuero**: Un material hecho a partir de pieles de animales tratados con un proceso especial.

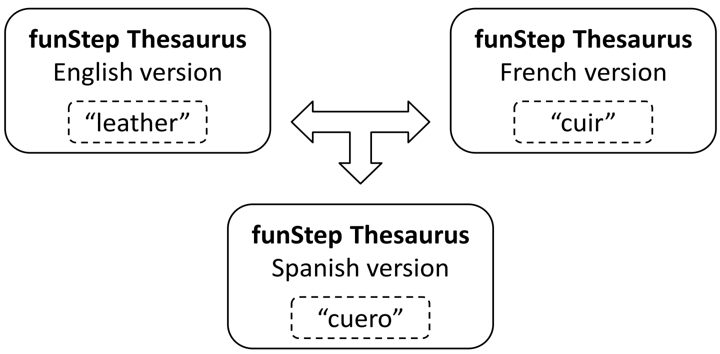


Figure 5.6 – funStep Thesaurus instances

The funStep AP236 Standard

To allow enterprise applications to interoperate seamlessly in information exchange, there is a need for a unified and standardized representation of product data [296][310], i.e., the funStep ISO AP236 standard [300]. AP236 defines the reference open structure for catalogue and product data representation under industrial domains of the furniture sector, helping on the information interoperability at a syntactical level. The External Classification; and Multilinguism modules of funstep standard are examples of relevance for semantic enrichment.

5.2.2 The funStep knowledge clients

The “funStep knowledge clients” are interfacing the engine with the funStep knowledge users, which can be administrators, customers, or stakeholders.

The role of funStep system administrators is mainly related to systems evolution and maintenance, i.e. KREs.

Traditionally, in the furniture sector, customer interfaces are mainly focused on simple product search. This architecture intends to go further ahead, enabling software developers with skills to develop enhanced intelligent products search engines based on knowledge reasoning.

The stakeholders’ interfaces are focused in providing the appropriate features to enable furniture stakeholders access to funStep knowledge in order to use it as reference and to contribute to their evolution.

5.2.3 Natural Language Dictionary (NLD) server

The NLD server complements the domain dictionary in the translation procedure of natural language information used within the AP236 standard data. For instance, the request for quotation sentence sent in the use case scenario was translated from “XPTO in leather” to the French “XPTO en cuir”. Despite the “leather” word has been translated by the domain dictionary, the “in” word was translated through a NLD.

5.2.4 Interoperability

The interoperability between the knowledge server and the knowledge clients (Figure 5.4), and between the knowledge server and natural language dictionary server is established and controlled by two sets of services, i.e., 1) knowledge services and 2) natural language dictionary services. These two sets provide the required services available through the Internet, that allows a global access and control of the knowledge functionalities of the funStep platform throughout its knowledge server, connecting the supporting functionalities of the natural language dictionary and, on the other side, the funStep knowledge clients themselves.

5.3 Semantic Enrichment for Intelligent Product Catalogues

The ISO 10303-236, i.e. AP236 or funStep standard, defines a formalized structure for catalogue and product data under the industrial domain of furniture sector. The standard is originally divided in six different implementation sets: catalogue, geometry, rules and expressions, room decoration, catalogue plus geometry plus expressions, and the whole standard [300].

The AP236 standard helps on the information interoperability at a syntactical level. In each implementation set there are several resources that establish the way how the information components must be represented. From all of the standard resources, three are of relevance for the standard semantic enrichment purpose, i.e. Internal Classification; External Classification; and Multilinguism [311].

Internal Classification

Each company has its own product nomenclature and structure. This is easily verified not only in the way catalogues are arranged, but also in the different designations companies use for the same concept. This way, normally legacy product taxonomies and classification systems, i.e. thesaurus or ontologies in the funStep knowledge case, are shared with other enterprises when doing business, in order to create a common understanding.

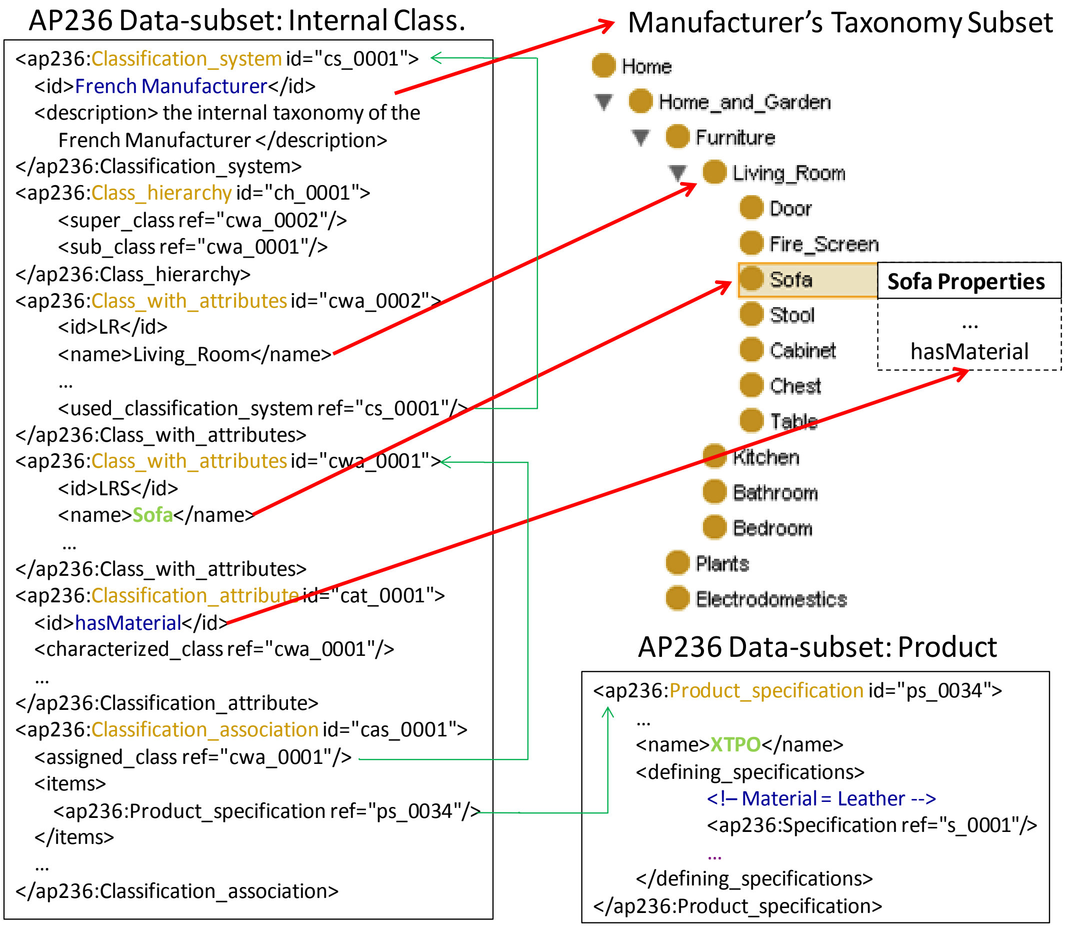


Figure 5.7 – AP236 Internal Classification Example

AP236 provides a mechanism to establish a direct link between the context and the classes of products present in the catalogue. The classification entities of the standard allow making one or more associations between classes of meaning and concepts, which may be helpful if one wants to make a taxonomy apart from its own just meant for the end user. With internal classification, the taxonomies are embedded in the product data file being exchanged among organizations.

Continuing with the "leather couch" scenario, if the user decides to go forward with the purchase of the "XPTO couch" in "leather", the manufacturer will have to produce it, and eventually will send the product details to the retailer, so that he can sell it in the future. At that time, the manufacturer can send the retailer its own taxonomy with the classified product. Figure 5.7 illustrates how that information would look in AP236 and how it matches the taxonomy terms (normally would be French due to the manufacturer nationality, but presented in English for easy-reading purposes).

The AP236 data model enables to specify a structure of classes, attributes and their relationships using the “class_with_attributes”, “class_attribute”, and “classification_hierarchy” entities. As illustrated in the figure, those classes also enable to establish a direct relationship with the product in the catalogue, i.e. “product_specification”, using the entity “classification_association”.

External Classification

For an improved business, networks of organizations may define, or use shared reference ontologies or thesaurus, instead of legacy taxonomies. In this case, when exchanging product information, they should classify their products using that reference nomenclature. AP236 provides a mechanism for that, i.e. the external classification.

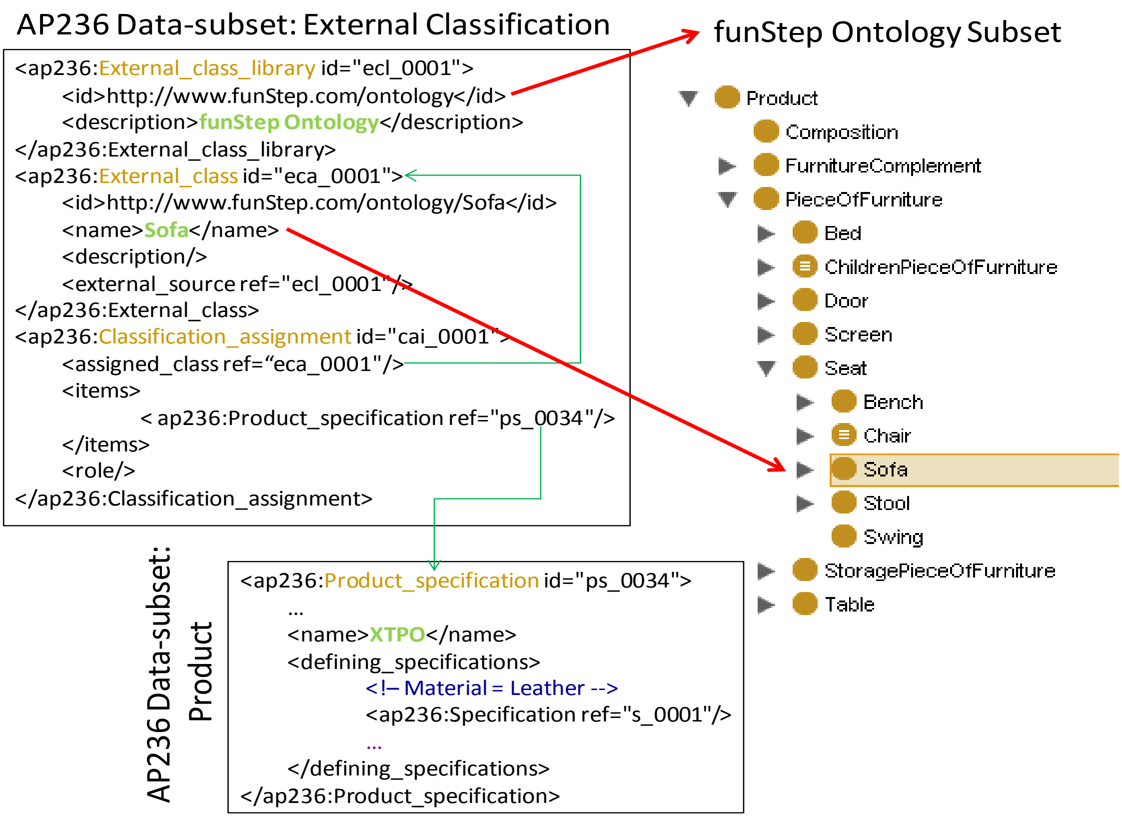


Figure 5.8 – AP236 External Classification Example

External classification enables a direct link between the context and the classes of products. Nevertheless, in this case, the mentioned link uses meanings that are expressed in libraries physically described in remote locations. Figure 5.8 illustrates how external classification works. The manufacturer links with the funStep reference ontology, instead of sending its own, using the AP236 entities: “external_class” and “external_class_library”, to identity the externally defined concepts, and “classification_assignment” to establish the link with the product in the catalogue.

To illustrate how external classification works, Figure 5.8 uses the same example of internal classification, however in this case the link will be established with the funStep reference ontology. Here, the AP236 entities used are: “external_class” and “external_class_library”, to identity the externally defined concepts, and “classification_assignment” to establish the link with the product in the catalogue.

Multilinguism Issues

For a better internationalization of the market, companies must be able to send their products and associated information in several languages, especially in the language the receiver company speaks. The multilinguism resource of AP236 addresses such need, so that any organization could receive data in their native language independently of where it was introduced or manufactured. In the “leather couch” scenario, multiliguism can be used at the time of catalogue publication or update (i.e., at manufacturer–retailer communication).

Multilinguism allows the translation of the attributes of AP236 entities in any number of languages. In Figure 5.9, it is possible to notice that the AP236 entity “language” is used to indicate the reference language. The entity “language_indication” points to the attribute that one wants to translate, and the “attribute_translation_assignment” specifies the translation itself. In this case, it is the value of the attribute “name” that is being translated from the French “cuir” to the English “leather”.

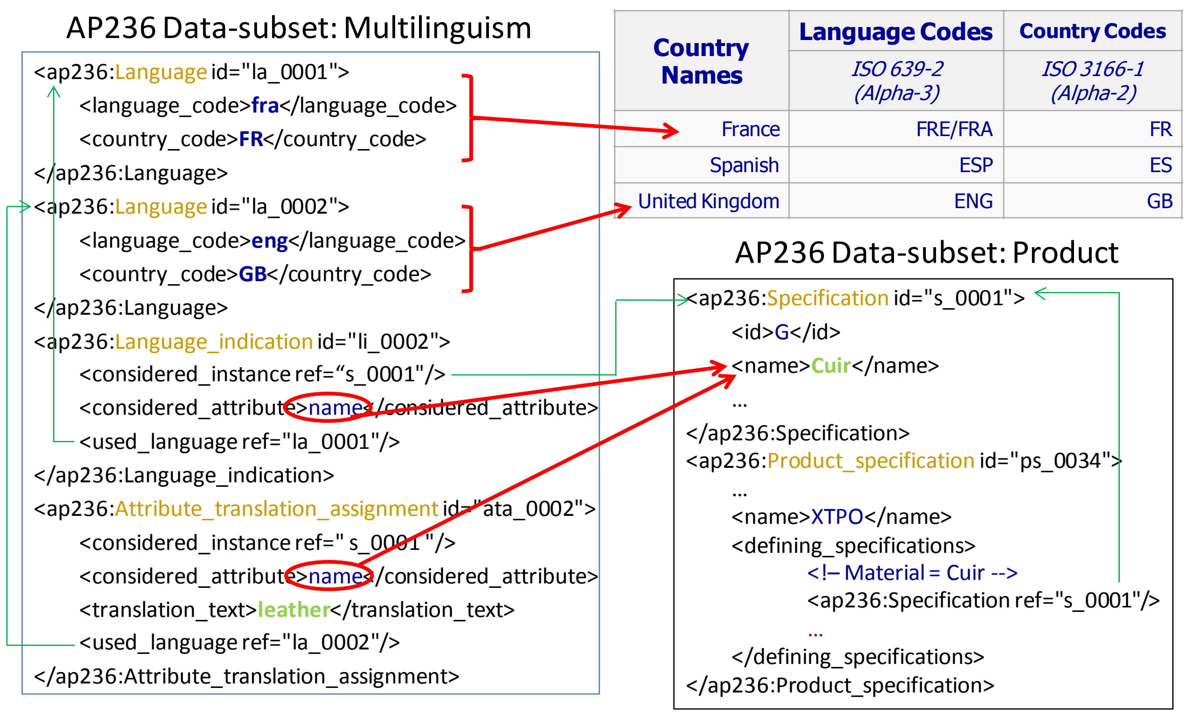


Figure 5.9 – AP236 Multilinguism Example

CADEF Application

However, there is already a tool (CADEF - CAtalog DEFinition) which is able to make semantic annotations. CADEF, as a configuration tool, allows the definition and management of all the possible specifications grouped by categories. Base in these categories and due the configuration process, the CADEF user could define the different variant a product could have. He is able to classify the products following the funStep ontology properties and as a result it is generated a catalogue in AP236 standard that establishes the links to the ontology in background, which are semantic enriched as presented before.

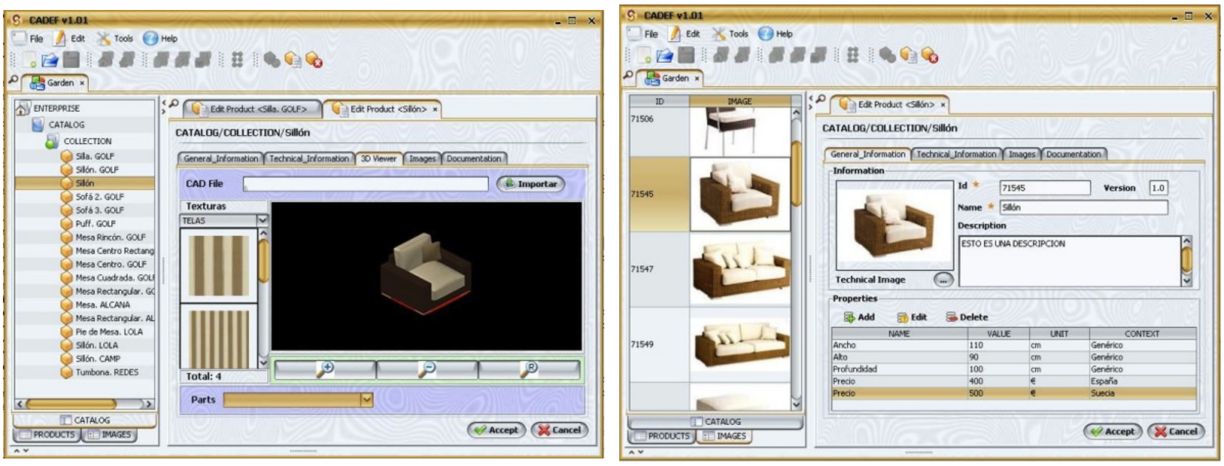


Figure 5.10 – AP236 Multilinguism Example

5.4 Management of funStep Knowledge Framework

It is the funStep managers' committee that maintains the KR elements. Having multilingual dictionary, thesaurus and ontology, there is a need to have such a domain expert managing each of the languages and KRE's used, that analyses the specific issues related to knowledge domain representation.

Following the "leather couch" scenario, there was a necessity of having the term "leather" translated in three languages: in English, French and Spanish. Since the word leather is a reference concept in the furniture domain there is needed to have it linked with all of reference terms in each of the languages (Figure 5.6). In this way the systems will be able to translate the reference terms to other reference terms avoiding misunderstanding in the meaning of such words. To have the reference concepts mapped with others in other languages, it is used prototype thesaurus a mapping tool for thesaurus concepts mapping establishment (Figure 5.11). The managers committee makes the action of establishing such mappings.

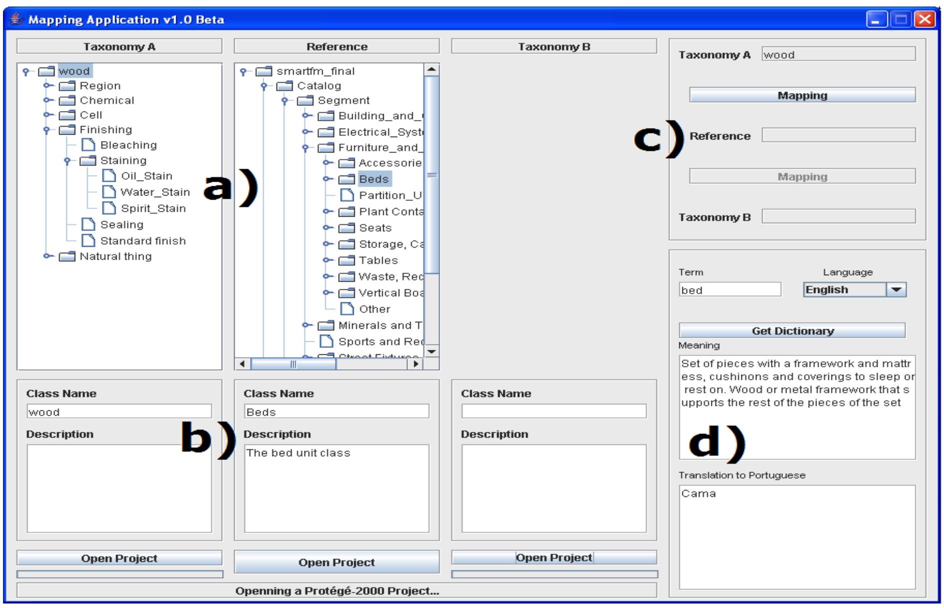


Figure 5.11 – Thesaurus Mapping Tool

This thesaurus mapping tool is organized in four areas (Figure 5.11) where in the area a) the user can select terms to map from the opened thesaurus and in the section b) appears the selected terms annotation. If the user identifies semantic similarities, then it can establish the mapping through the available functionalities in area c). In the area d) it is facilitated the access to the furniture funStep dictionary for human check and validation.

To relate all the KRE concepts it is used MO following the tuple-based mappings as defined at sub-chapter 2.3.2. It is foreseen by the author, the integration of this thesaurus mapping tool functionalities with the mapping tool presented at sub-chapter 4.5.2 and accomplish it with MENTOR for ontology evolution. The objective is to have at the end a MENTOR tool prepared for the multilinguism and able to make semantic enrichment of product standards (complaint with the funStep knowledge architecture).

5.5 Conclusions

The proposed funStep knowledge architecture provides enterprise and manufacturing systems a semantically seamless communication with other stakeholders up and down the supply chain.

The development of common international standards and methods helps to master the global value chain. However, alone, product data standards do not solve today the EI problems. Indeed, typically each stakeholder has its own nomenclature and associated meaning for its business products. Therefore, organizations from similar business environments are having trouble cooperating.

The furniture sector is no exception, and being a manufacturing sector based mostly on small and medium enterprises is having the problem proliferated by millions of organizations. For instance, a group of enterprises that share product data information in their business activities need to have a common semantics to understand each other. Otherwise their systems might understand the data structure but not their meaning.

The author, under the funStep initiative which is continuing the activities begun by the IMS SMART-fm project, proposed the semantic enrichment of standardized product data as a solution for making interoperable intelligent manufacturing systems a reality. He endeavours to gather the tacit knowledge that furniture domain stakeholder's hold into machine interpretable KBs which should be stored in a structured organized way, where syntax and lexical semantics are integrated (explicit knowledge).

This allows enterprises to keep their internal terminologies and classification systems, and still remain interoperable with their business partners, through the usage of knowledge mapping procedures. The funStep knowledge architecture uses different KREs as catalysts to enable such semantic interoperability. Together, the domain dictionary, the thesaurus, the reference ontology and the AP236 standard itself act as explicit knowledge repository and reference lexicon for the application domain.

To complement, the domain dictionary in the translation procedure, a natural language dictionary is used. And since knowledge is not constant, the architecture also considers its maintenance by a group of administrators, as well as the interaction with the users, e.g., customers, and stakeholders.

6 oLEARCH – Ontologies LEArn by seaRCHing

Due to the worldwide diversity of communities, a high number of ontologies representing the same segment of reality that are not semantically coincident have appeared. To solve this problem, a possible solution is to use a reference ontology to be the intermediary in the communications between the community enterprises and to outside. Since semantic mappings between enterprise‘s ontologies are established, this solution allows each of the enterprises to keep internally its own ontology and semantics unchanged. However information systems are not static, thus established mappings become obsoletes with time. This chapter presents a suitable method that combines semantic mappings with user‘s feedback, providing an automatic learning to ontologies enabling auto-adaptability and dynamism to the semantics of the information systems.

The feedback given by the users are used to get some patterns, which through appropriate analysis and machine learning algorithms also in association to statistics enable ontologies to learn, readjust or updated their concepts and consequently to be improved.

As mentioned by Brewster in [312], the evaluation of the output of OL systems remains a major challenge. The usual approaches to ontology evaluation have largely been based on quality control of the ontology building process (OL) and ensuring the ontology abides by certain principles [313][314]. But, on this case it is intended to provide another kind of OL, which intends to learn from its users to constantly improve the semantics interoperability between its representations systems maintaining its knowledge updated. Knowledge maintenance is then a facilitator to the dynamic information systems interoperability and consequently is the main actor to the OL process.

Author found appropriated to define “oLEARCH - Ontologies LEArn by seaRCHing”, as a new concept related to ontologies able to change/adapt their knowledge (to learn) through their users’ patterns of searching/reasoning. The concept was inspired from the concept LEARCH defined by Ratliff et al. in [315] that means “LEArning to seaRCH” and was defined to represent algorithms for imitation learning in robotics with the main purpose to search something.

oLEARCH is a products search engine application available to users by Internet. This system learns from user’s searched products concepts improving the KB. This KB is consisted by at least two ontologies, the MO and the reference ontology.

The oLEARCH function uses an algorithm supported in an instance-based learning approach based on user interactions. In instance-based learning, training examples are stored verbatim, and a distance function is used to determine which member of the training set is closest to an unknown test instance [197]. In oLEARCH, such distance function is represented by the *semantic distance*, which is the inverse of the *semantic relatedness* between the users introduced concepts and the products classified in the reference ontology. Thus, oLEARCH provides to the users a set of products that are close to their introduced concepts in terms of *semantic relatedness.* Then, users are able to select the most appropriated product from this set of possible choices. These last users’ selections are also used, as a last feedback, to increase the *semantic relatedness* weight of the selected products associated concepts.

6.1 *oLEARCH Implementation Architecture*

The oLearch implementation architecture (Figure 6.1) is composed by a reference ontology, a MO and two GUIs: the Web and the Administrator GUI, supported by a set of webservices.

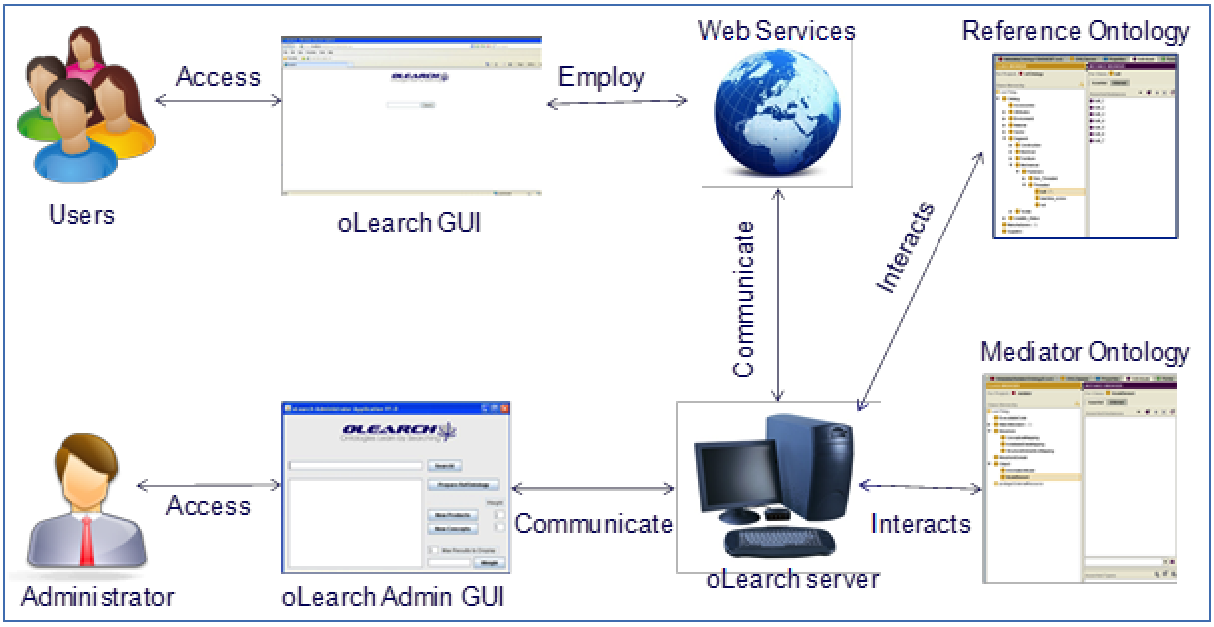


Figure 6.1 – oLearch Architecture

6.1.1 **oLEARCH Web GUI**

The oLEARCH Web GUI (Figure 6.11) is a search engine application built in JSP technology that uses two specific web services able to learn from web users. It is similar to Google[38]. The main difference is that oLEARCH uses ontologies as KB, and it learns new concepts relating them to previous existent ones. The Web GUI is the main oLEARCH application output; its main function is to search for products. When the "SEARCH" button is pressed oLEARCH will compare the entered concepts with the KB. If there is any product related to such concepts, they appear on the screen as Google does. If not a message saying that nothing was found appears. Consequently, when a result product is selected its related concept weight or importance is increased automatically; also as Google does with web pages links. Next time such selected products if searched will appear on the top of the results.

6.1.2 **Reference Ontology**

Ontologies represent knowledge through the use of elements with its properties, rules, facts and relations. OLEARCH needs at least one ontology to represent the knowledge of a domain, with instance products classified on it.

[38] Google Search (or Google Web Search) is a web search engine owned by Google Inc. Google Search is the most-used search engine on the World Wide Web, receiving several hundred million queries each day through its various services [316].

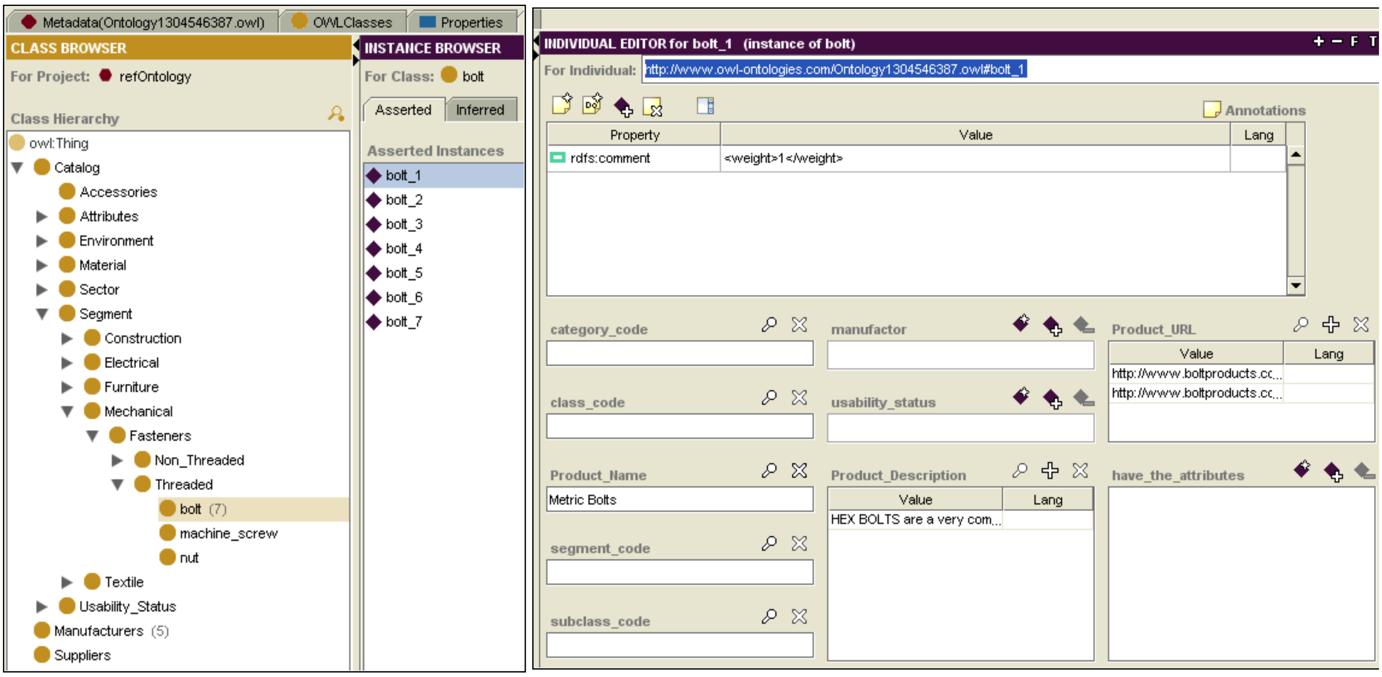


Figure 6.2: Reference Ontology classes and instances.

Figure 6.2 shows an example of a bolt product existent in the reference ontology. This example represents a “Metric bolts”, which for at the time it had a weight of one. All the concepts used in this ontology have a respective weight to measure its importance, representing its appearance frequency in users searching processes. This weight is applied into the reference ontology’s classes and instances in rdfs comment property, to avoid adaptations inconsistencies. Reference ontology after a first search will have its own weights represented in the MO.

6.1.3 **Mediator Ontology**

In oLEARCH it is used to record mappings between “new concepts” and/or between existing reference ontology concepts and obviously to manage their weights. Such mentioned mappings are of the conceptual type.

When a user searches for more than one concept, oLEARCH assumes that they are somehow connected and so it creates relations between them and records it in the MO. Figure 6.3 contains an example of how a relation is stored into MO. In this case is a relation between a “bolt” and a “parafuso” concept. Looking carefully to the example it is possible to see that such conceptual mapping has already a weight of four. This means that such concepts were searched together already four times, or after introduced together once the concept that is not present at a reference ontology was already searched three times. MO also records elements that were searched alones. Figure 6.3 shows a case of this type.

6.1.4 **Admin GUI**

The oLEARCH Administrator GUI (Figure 6.16), is the administrator tool where it is possible to set some parameters such as the number of results to be displayed on the regular oLEARCH GUI, set the minimum value for a searching concept to be considered as a possible “new product” or as a possible “new concept” and also to increase manually a concept weight. This tool has a display to show this possible new products and concepts. This application also provides the possibility to search for an element, by pressing “Search” button.

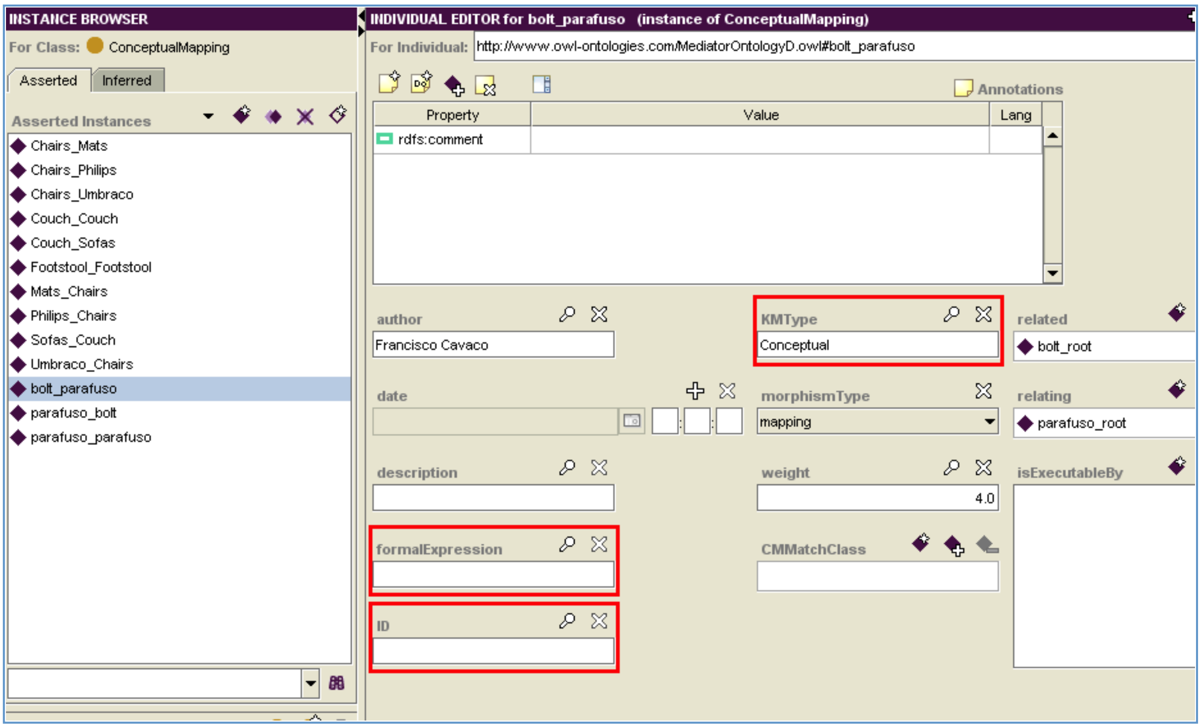


Figure 6.3: Mediator Instance representing a relation between ‘bolt’ and ‘parafuso’.

There is other administrator functionality related to the button called “Prepare RefOntology” which prepares the reference ontology to be used in oLEARCH. Basically, it adds a property called “weight”, with value equal to one, to all classes and instances of the ontology. This weight is added in the comment property as shown in Figure 6.2. Also, in the mediator ontology it will create two new Information Models related to the chosen reference ontology data (PatternsInfoModel and refOntology shown in Figure 6.4), to prepare MO to receive information about searched terms,

6.1.5 OLearch Web Services and Functions

There are two oLEARCH services, one to get products and another one to update a product weight. These services are available to final users through web services and locally to the administrator.

Get products service

This service is activated when a user searches for a single or multiple concepts. This service is represented by two main functions, one to take care of the searching concepts and another to filter and return the related products found. The way a concept is handled depends on if it is a class, instance or a property, in the KB.

Increase weight service

After the searched products are returned to the user, he might select the desired one. This selection activates the second service which is used to increment the product weight that he had chosen. It searches for the concept in the reference ontology and increases its weight. With this action, the selected product importance will grow which makes it appear first in future searches. This option can also be manually used in the oLEARCH administrator tool for testing purposes.

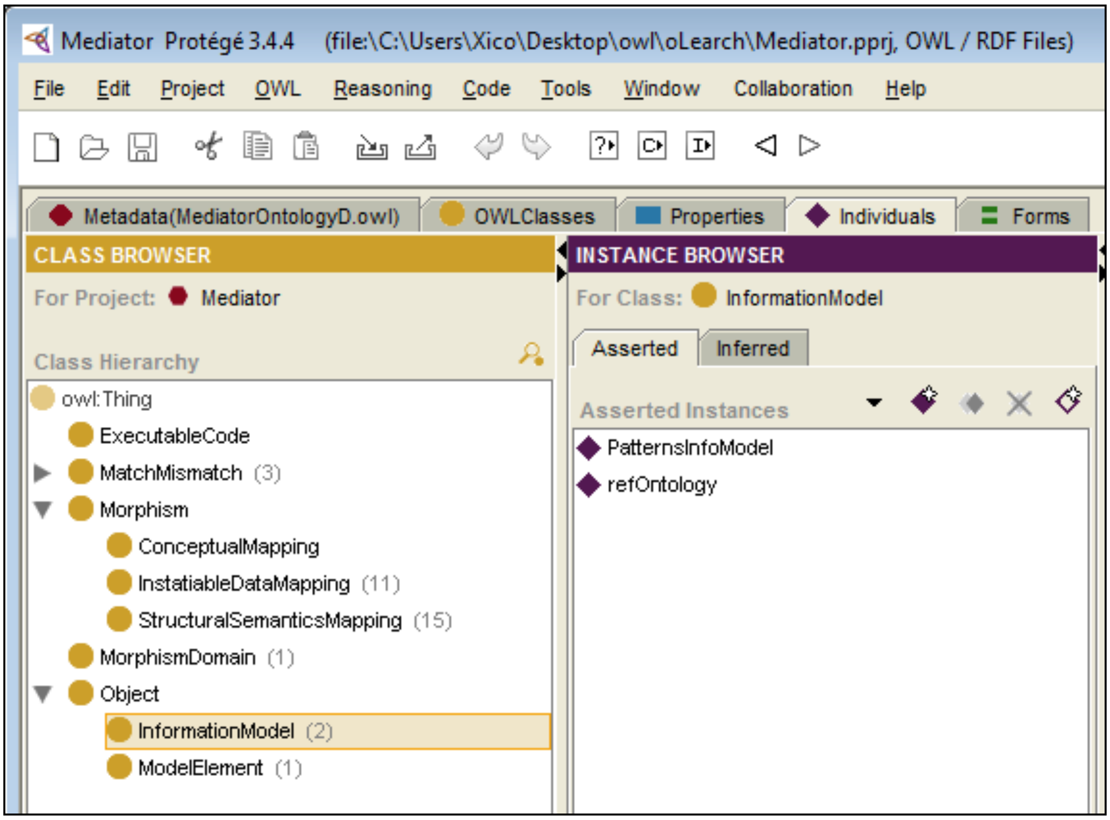


Figure 6.4: Information Models in Mediator.

There are two specific admin functions, which from specific search patterns define the existence of possible “new products” or “new concepts”.

New products

Possible new products are the concepts searched by users which don’t exist in the reference ontology and which were not related to existing concepts. This will search in MEDIATOR for concepts that don’t exist in the reference ontology and that were searched more than the minimum settled in the Administrator GUI.

If this value is defined for two, this means that when a word is searched more than two times without semantic relation to an existing one, it is probably a new product.

New concepts

New concepts are the concepts searched by users, which have a semantic equivalent concept in the reference ontology, but are not the same. Using Figure 6.10 example, “Couch” has a relation with “Sofas”, if it is more used than “sofas”, “couch” will become a new concept for sofas. OLEARCH implement this by getting all existing relations where one of the concepts doesn’t exist in the reference ontology.

If administrator defines one as a threshold for this functionality, it means that for a concept be considered as a new concept, it just needs to be searched more than one time, or have its weight bigger by one, in the relation to the related concept of the reference ontology.

6.2 oLEARCH use cases

6.2.1 oLEARCH use from technical perspective

The first thing oLEARCH does with the searching concepts is to check if semantic relations are registered between them in the MO. If yes oLEARCH increases its weight. If no relation exists between these concepts, then oLEARCH creates them. These relations between two model elements are created as morphism of conceptual type. Then, if there are relations to reference ontology concepts the service has to select its products returning them to the user ordered by relevance (products with higher weights will come first). It's in this step where the referred global lists are used, starting to remove all child classes existing in the class list to avoid having duplicate results. For each concept (node) a label is created with ID, name, type (Classe, Instance or MO element) and URI. Also, graphs are created, with origin, destiny, weight and if it is or not an instance (product) properties. There are three kinds of global lists, one with reference instances, another one with reference classes and a last one with MO model elements. The logic here is if it's an instance (product) then it will be connected from "_inic" (origin) to "_out" (destiny). If it's a class, then the algorithm will search for its child classes and instances creating all the possible connections. Regarding MO elements, these will be separated as if they are related to a reference class or instance and then handled accordingly. Only instances are connected to the destiny "_out". The Figure 6.5 presents a possible graph with some mentioned connections.

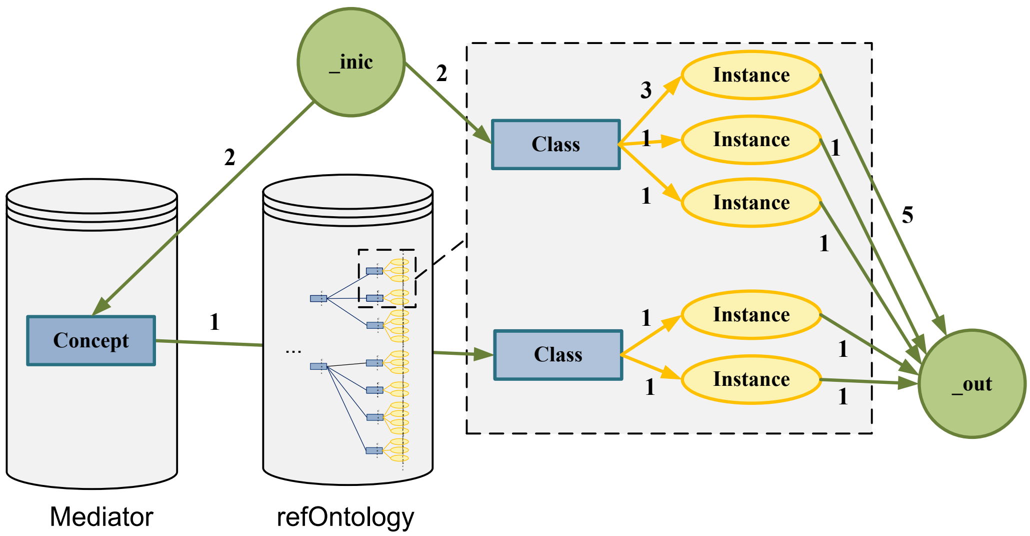


Figure 6.5: Graph representation

Note that properties are not represented because, as mentioned before, oLEARCH returns their classes instead. It is also possible to see that only instances can be connected to element "_out". This is because oLEARCH will only return instances (products).

The returned results are presented by its weight order, using Dijkstra algorithm for this decision. Usually Dijkstra is used in graphs to find the closest way. What oLEARCH does is to invert weights in order to have the biggest one with lower value. This will make the higher weights will come first than lower ones. This weight's transformation in distance basically consists in inverting all existing weights using the following formula:

$$\text{newWeight} = \text{maxWeight} + 1 - \text{oldWeight}$$

Figure 6.6 demonstrates the resulting scenario after this weight transformation be applied to the scenario shown in Figure 6.5.

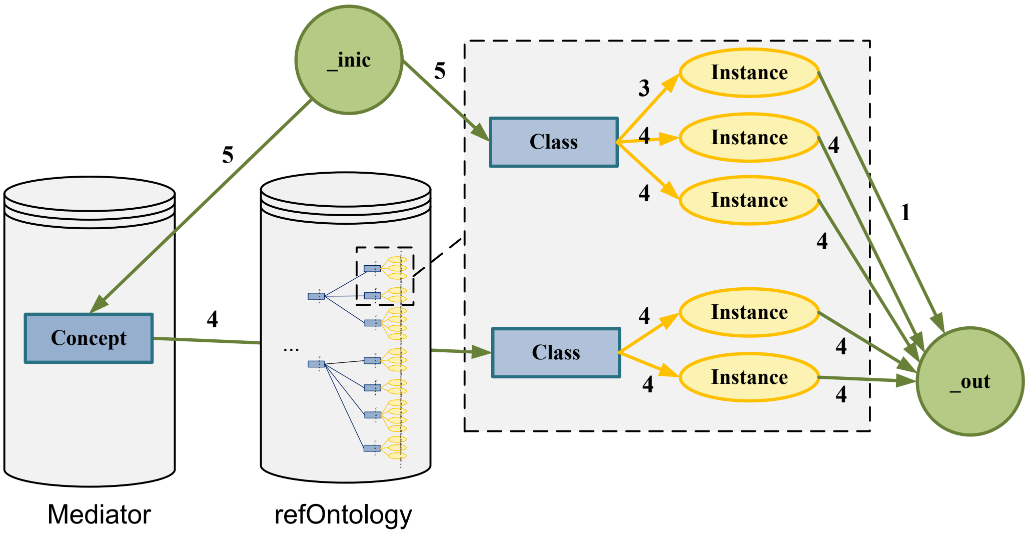


Figure 6.6: Graph representation after weight transformation

It is now possible to apply Dijkstra algorithm to the existing graph. It will be applied as many times as the value of number of results is defined in the administrator GUI, returning a different product each time.

Next figures represent a practical example of the explained theory used in oLEARCH. Figure 6.7 presents parts of the connections established when it is searched for “Chairs” in the initial oLEARCH state.

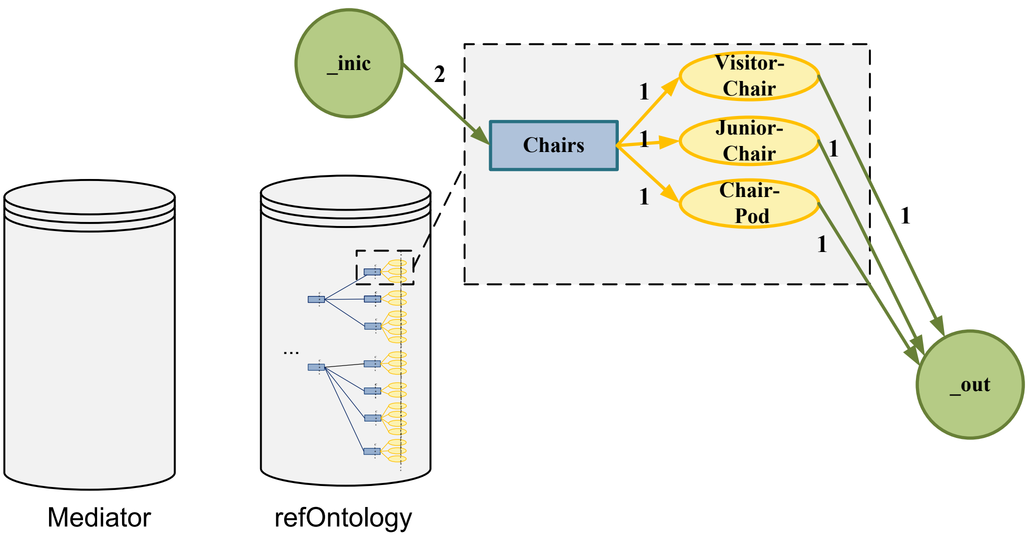


Figure 6.7: Graph created when user searched for “Chairs”.

Then, Figure 6.8 shows graph result when the user searches for “Couch”, which doesn’t exist in the KB.

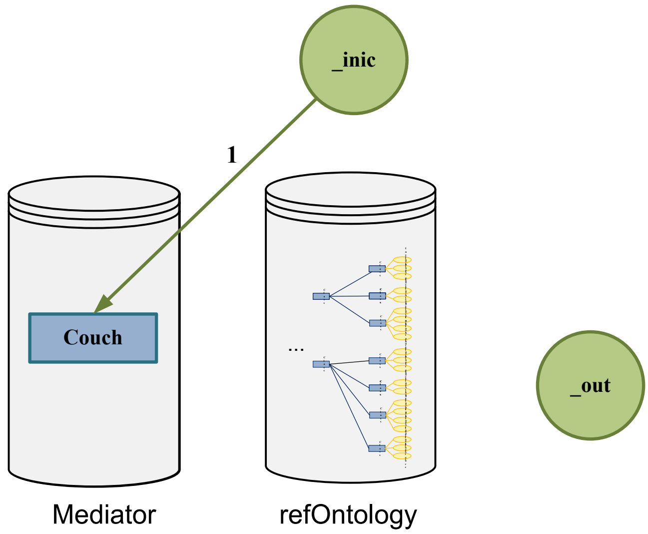


Figure 6.8: Graph created when user searched for "Couch".

Finally, after the user searches for "Couch + Sofas" it is created a relation between such concepts. Figure 6.9 scenario represents a "Couch + Chairs" search. It is possible not only to see the new relations that are created but also the respective weights increasing.

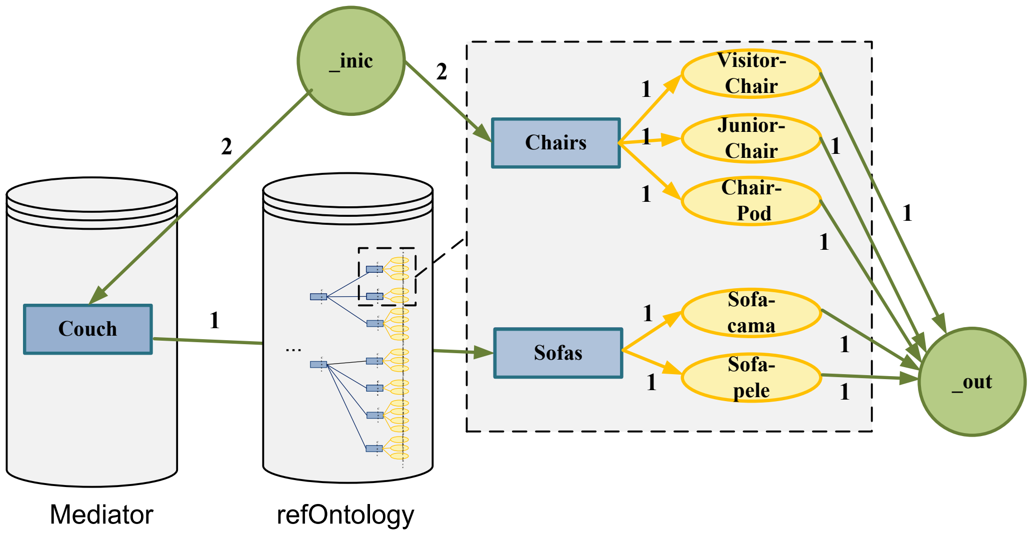


Figure 6.9: Graph created when user searched for "Couch + Chairs".

After this process the weight transformation in distances is applied and the result is displayed in Figure 6.10.

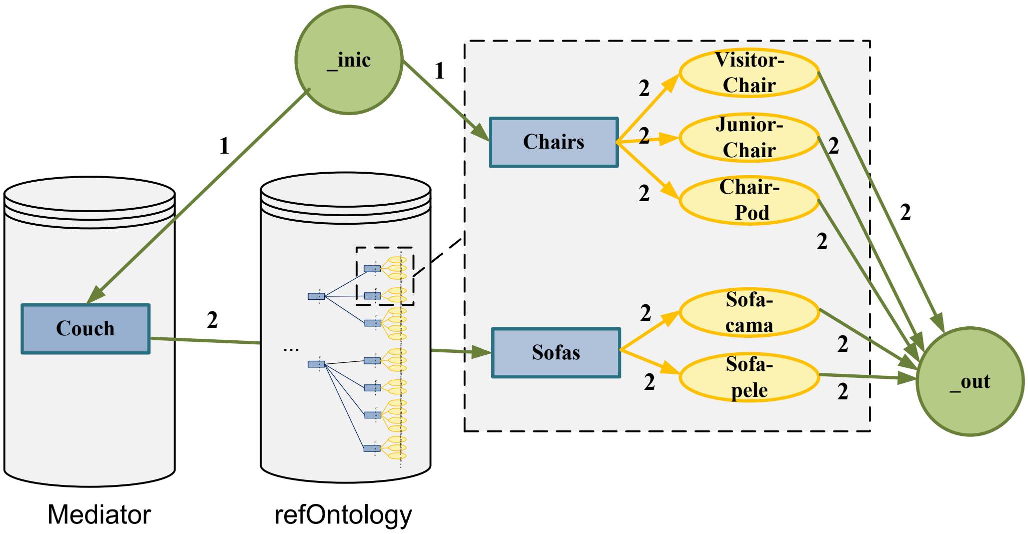


Figure 6.10: Graph after weight transformation.

After transformation Dijkstra algorithm is applied, returning all the products, till maximum products settings, starting first from the one with lower weight. From the example of Figure 6.10 the first products returning are the chairs because they all have weight sum, from "_inic" till "_out", equal to 5 (less than Sofas which have weight sum equal to 7).

After this process, if the user selects one of the searched results then oLEARCH will communicate back to the reference ontology the user choice in order to increment such selected product concept weight.

Figure 6.11: Search results for keyword "Chairs".

6.2.2 oLEARCH use from users perspective

OLEARCH use demonstration starts by searching the concept “Chairs”. This search returns all the Chairs existing in the Ontology (Figure 6.11).

Then, if the search concepts are “Chairs+Mats” which returns all Chairs and Mats in the Ontology, as presented in Figure 6.12. The reason why chairs are displayed before the mats it is due to last search, which had increase “Chairs” weight. In the Mediator, we have now a new relation between “Chairs” and “Mats”.

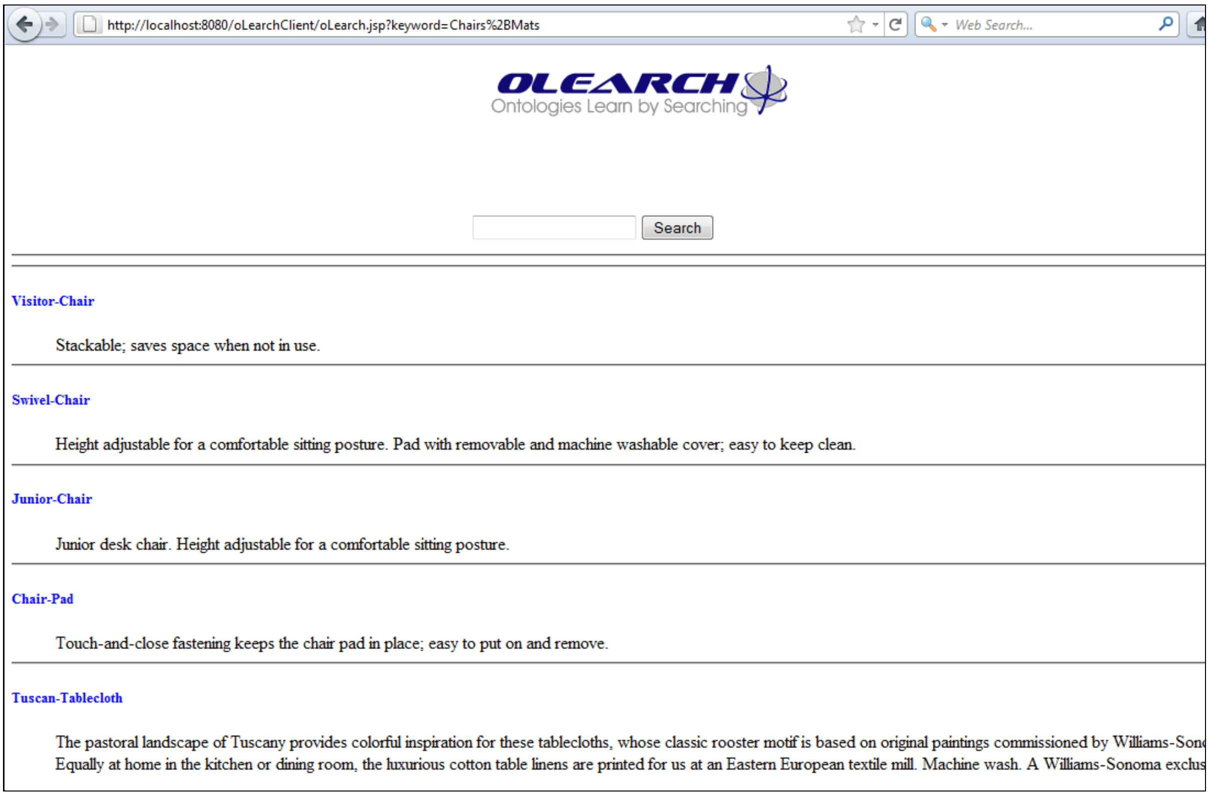


Figure 6.12: Search results for keywords “Chairs+Mats”.

Repeating the search using the concept “Chairs”, the result is the expected one that oLEARCH returned in Figure 6.13, not only chairs but also mats.

This result is due to the relation that was created in the mediator, by the previous search, where “Chairs” was connected to “Mats”. In this scenario if “Mats” gains importance (weight) till it’s higher than “Chairs”, then it will appear first in the search result, even if the user searches only for “Chairs”.

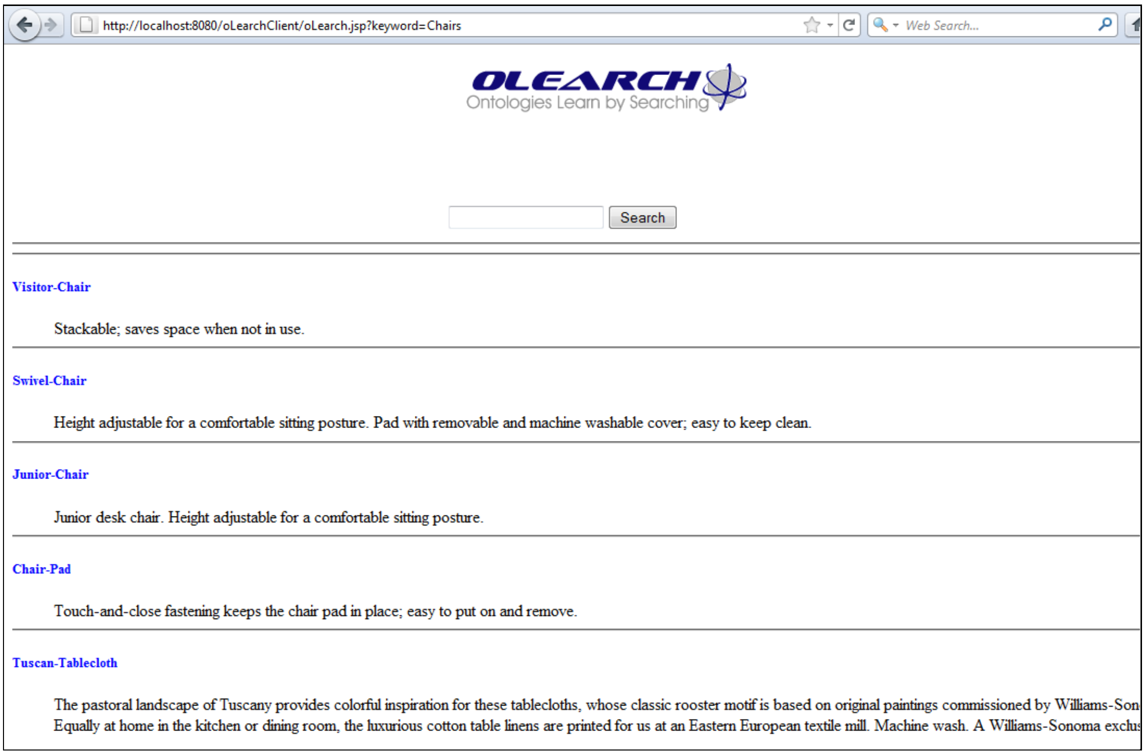


Figure 6.13: Search results for keyword "Chairs".

Figure 6.14 and Figure 6.15 show two extra tests. The first one is that searching, for example, "Mats" and/or for "Chair-Pad" several times will increase its weight. This weight increase will cause a change position when displaying the results. So searching again for "Chairs" the result will have a chair-pad product before some chairs.

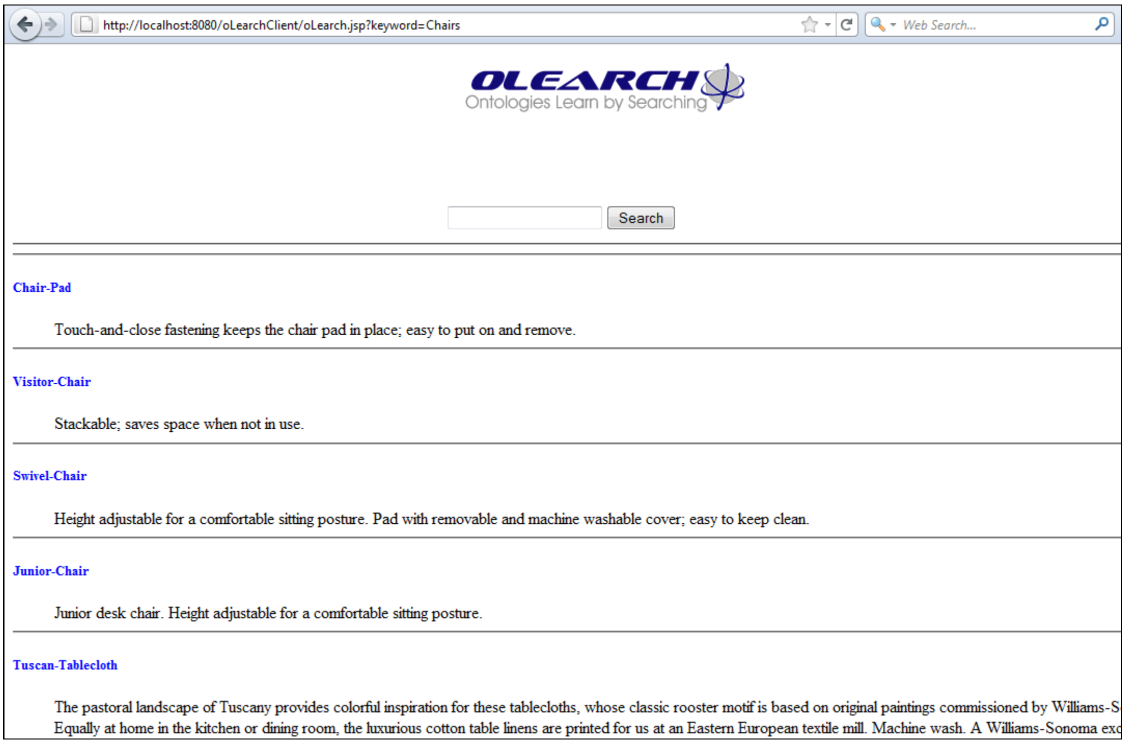


Figure 6.14: Search results for keyword "Chair-Pad" after searched several times.

Clicking on the product "Junior-chair" will lead to the second extra test, which shows that searching again for the same keyword, the product "Junior-chair" is shown in a high position than it was before.

This is presented in Figure 6.15.

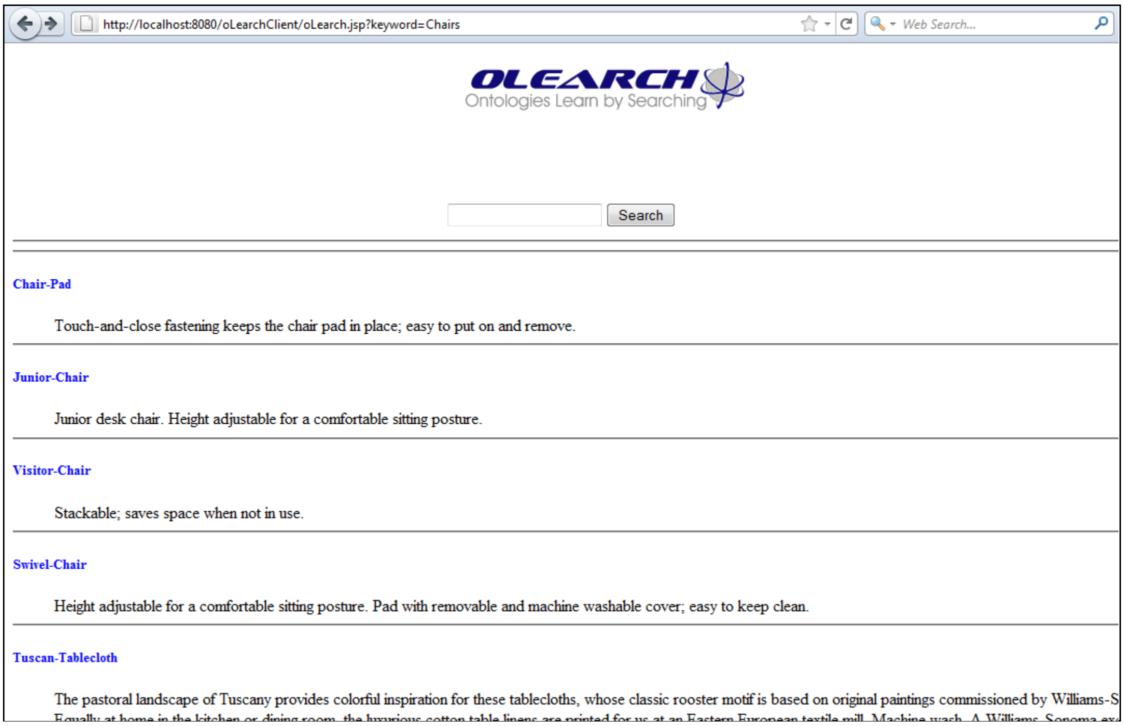


Figure 6.15: Search result for keyword "Junior-chair" after selected from the last search result.

This happens because of the searching frequency. It could be also explained as the following. Imagine that the weight of product A is three, weight of product B is two and weight of product C is one. They are all related to each other. Searching for A, the result will be A then B then C. After this search, product A has weight equal to four. If product C is searched two times, its weight will be equal to three. Now, when searching again for product A, the result will be A (with weight equal to five), C (with weight equal to three) and B (with weight equal to two).

6.2.3 **New concepts Functionality**

Considering that a client not familiar with the web service starts browsing the furniture domain database, using terms not existent in reference ontology. Those terms may be considered either a mismatch or a different jargon that such client habitually uses. The track of such behaviour may induce the system to react ignoring repeatedly the entry or wisely understand that the term is new for the reference ontology, consequently it could be the new term to be used in the future. This case represents time and money saved since it avoids domain engineers meet constantly with knowledge engineers to discuss enterprise's knowledge update as for new terms inclusion in the reference ontology. With this procedure, errors will be also reduced since the system could dynamically learn that a new term introduced by users could initially be considered similar/equal in semantics to an existent one as demonstrated before.

The used reference ontology has some sofas, although if a search concept is "Couch", oLEARCH won't return any product. Searching for "Couch+Sofas" it will return all the sofas. Then if other users search for "couch" more than for "sofas" will result in a "couch" weight higher than "sofas" weight. Thus, if the "new concepts" threshold is set to five, "couch" will be considered a new concept if it has a positive difference to "sofas" related weight of at least 6.

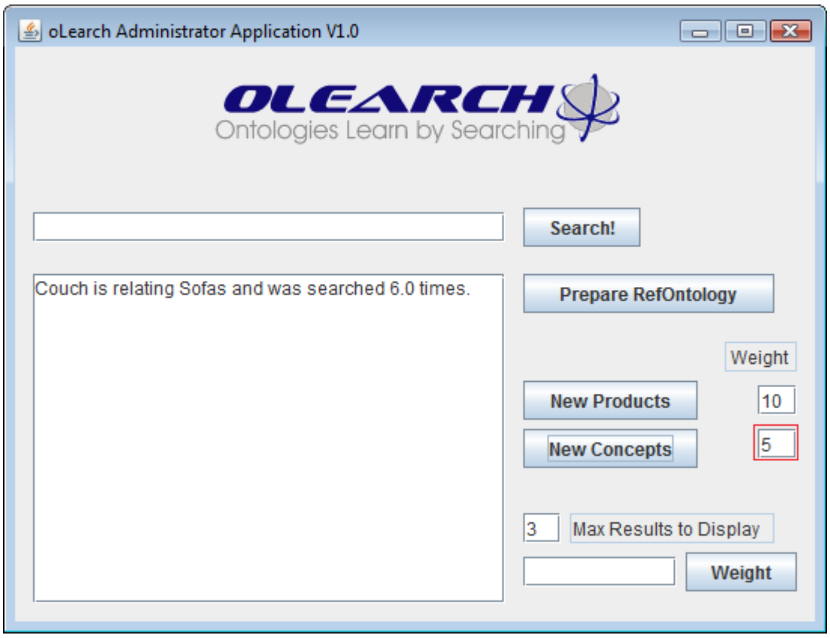


Figure 6.16: oLEARCH Administration Application result for "New Concepts" button.

6.2.4 New products functionality

Finally, another scenario is when the searching concept for an inexistent product, moreover with combined specifications or others details. It is expected that the system intelligently react and suggest MENTOR managers that some kind of new product has been requested (and design teams should be aware). In this case, the added value is more related to the time saved, because if there is any new product on the market, enterprises aim to react promptly. This system is able to warn design teams for a possible new product appetence, anticipating precious market information to the company.

For this test scenario, 'New Products' weight is set to 10, as shown in last picture. This means that if a concept is searched more than ten times and it doesn't exist on the ontology it is probably a new product. If it is simply searched for it ten times and it will return it as a possible new product. Thus if an administrator click on the "New Products" button after the mentioned steps, will appear "Couch as a possible "New Product" (Figure 6.17).

6.3 oLEARCH use in a business network supported by multi-agent systems

When working in collaboration environments enterprises need to communicate and understand each other's information automatically. However, sometimes it is not possible due to some interoperability issues. After investing in interoperable solutions, organisations can no longer support or accept these issues, because this means a loss in time and money, and sometimes maybe products and items. It is, therefore, important to maintain a sustainable interoperability of enterprise systems, whether they are communicating via P2P exchanges, standard based or even using reference ontologies. For this reason, it was implemented a framework based on CAS (Complex Adaptive System) namely the CAS-SIF (CAS-based sustainable interoperability framework) proposed by Jardim-Goncalves et al. in [317], which in this case are focused on the validation section that is liable for monitoring and proposing a (re) integration of organisations. In this line, Wooldridge in [318] says that decentralised multi-agents are considered to be an added value in the monitoring services implementation, assuring organisations' independence. Following this observation, this work uses MAS.

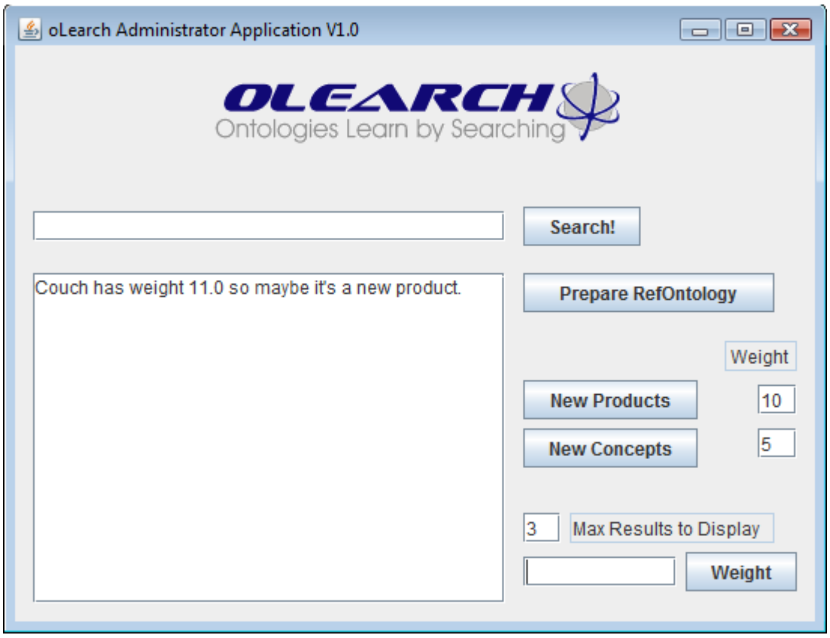


Figure 6.17: oLEARCH Administration Application result for "New Products" button.

6.3.1 Multi-Agent Systems (MAS)

There are several definitions to describe an agent but most supported terms are described in Wooldridge and Jennings ([319]) where they identify agent's properties: Autonomy - agents make decisions without human intervention; Social ability - agents interact with other agents via an agent communication language; Reactivity - agents react to changes in their environment; Pro-Activeness - agents have their own goals and besides reacting, they are capable of initiative.

Taking in consideration these descriptions, in particular the features of social and reactivity, it is common that the agent is capable of interacting with other agents, humans, or with the surrounding environment. This brings something new to the software technologies, i.e., communication and teamwork between software, in this case, between agents, and this is called Multi-Agent System [320].

The great advantage on using MAS for the implementation of CAS-SIF is that they are capable of cooperation, collaboration, negotiation, etc., and they understand each other via an agent communication language based on speech act [320], thus avoiding agent's interoperability issues. Therefore, due to these features, MAS is being used in different areas, from industrial applications, to telecommunication and multi-robotic systems.

6.3.2 MIRAI Framework

The role of MIRAI (Monitoring morphIsms to suppoRt sustAinable Interoperability of enterprise systems) framework is to monitor all the mappings that exist among the several models used by business partners in the same collaborative network (Figure 6.18), controlling the changes, warning and proposing new mappings, thus preventing interoperability problems that could cause a destabilization of the network harmony. Enterprises' privacy is assured since each one has its own MIRAI associated to an internal CM that tracks the morphisms it maintains with is direct partners.

Interoperability is often guaranteed by setting P2P mapping morphisms with different partners, or in optimized networks using neutral data standards to regulate communications (section 2.1). MIRAI was initially proposed by Ferreira et al. in [321] with the objective of monitoring the existing mappings and model versioning's stored in each enterprise's CM and timely detect the changes in the morphisms, proposing to the user a possible solution and preventing significant transient periods where interoperability in the network is not assured. The detection is carried as soon as CM changes, triggering an agent to search model differences.

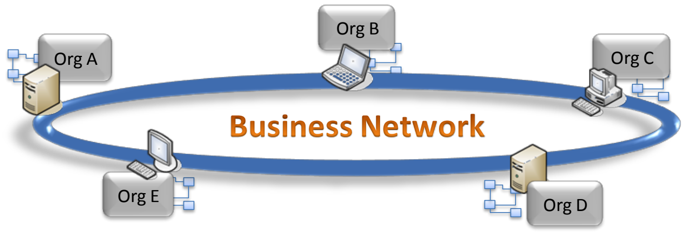


Figure 6.18: Business Network

Indeed, when a versioning on one of the models is detected, MIRAI triggers a warning and automatically proposes new mapping morphisms to the user realising the integration intelligence layer of CAS-SIF. This new suggestion is based on the 5-tuple mapping expression explained before. Within this framework, a new mapping is created to respond to the evolution of an organisation's IS (storyline: A→B exists; A→A' evolution; A'→B mappings proposal). However, not always the first mapping proposal is the one the user prefers, thus the need for the system to learn from experience is required, adjusting more and more to the user's preferences.

The extended MIRAI architecture is represented in Fig. 4 already including a learning module to respond to the above. It is divided in five blocks:

1. Intelligent Supervisor to detect the harmonization breaks by performing a scan in the organization's CM in search for new morphisms or morphisms evolutions. This agents' block checks if the changes occurred have an impact in the system's interoperability status, and if it does, it proposes a possible solution;

2. External Communicator which is responsible for the communication between different MIRAI's of a collaborative network;

3. Administrator GUI is the interface between the MIRAI and the human user. In the future, it will be further developed into enclosing the decision support services envisaged in CAS-SIF;

4. Life Cycle Monitor which is a maintenance block to check if any agents died, and is responsible to resurrect them if required:

5. And finally the oLEARCH block which encloses the ontology learning algorithm previously described and is able to do make prognosis. The User's GUI is also a part of this block since as envisaged in the algorithm; oLEARCH requires knowledge gathered from the user

interaction.

6. Following this idea, an extension of the MIRAI architecture capabilities to a dynamic level is obtained with the implementation of the oLEARCH. This brings to MIRAI, knowledge from the users feedback that will be used for ontology learning. This learning will help in the prognosis proposals for interoperability re-establishment, which will make the proposals more accurate and the human intervention less used.

Because of the new architecture characteristics and the features of MAS, MIRAI's adoption of ontology learning is called on this case multi-agent learning. The learning agents cooperate to share the learning lessons, so they are able to locate and translate disparately referenced concepts and improve from others [169]. Decentralised multi-agents are considered to be an added value in the monitoring services, thus it was decided that MIRAI would be composed with a seven agents MAS divided by the five blocks, six of those agents belong to the original MIRAI [321]: Agent Monitor Mediator and Agent Momo belong to the Intelligent Supervisor; Agent User to the Administrator GUI; Agents Persistor and Agent Persistor Police to the Life Cycle Monitor; and finally the Agent Communicator the External Communicator. To the extended MIRAI was added a new agent (i.e. Agent Pattern Learning) that does the monitoring of new semantic patterns created with the introduction of new concepts using the services of the oLEARCH.

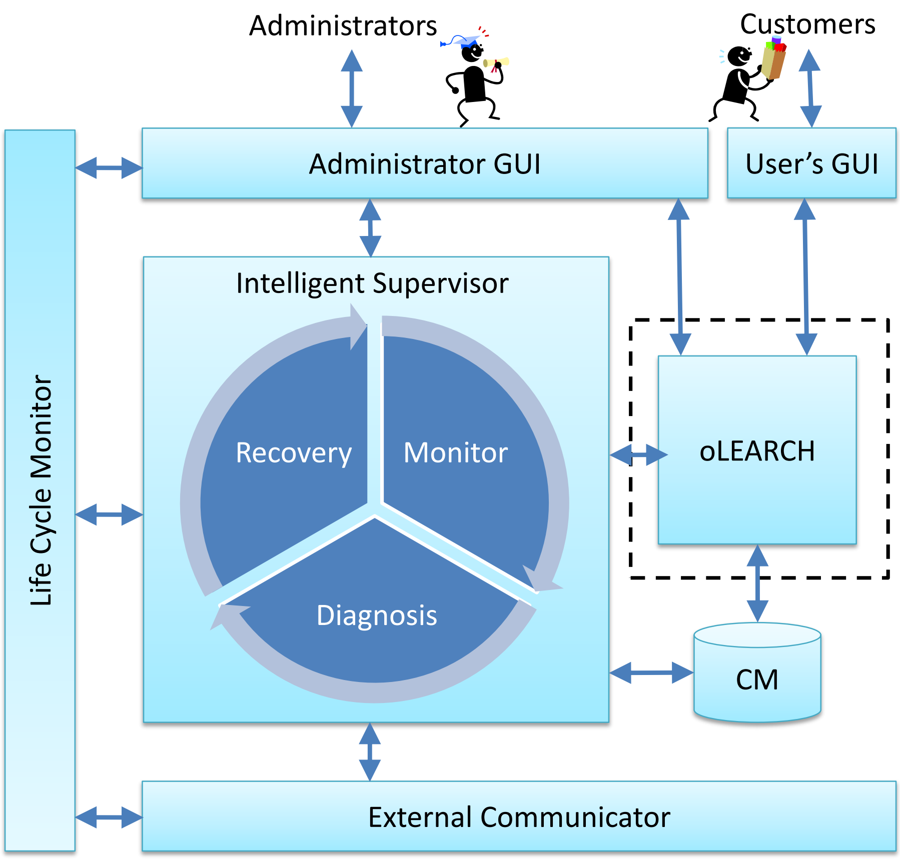


Figure 6.19: Architecture.

The oLEARCH technology complements that and creates an understanding environment between the organisations inside the network. Knowledge maintenance will assume a key role to dynamically update reference ontology, since enterprises concepts are changing fast in today's competitive market. The possibility of the system to learn from the user interactions and understand each other without the constant assistance of the decision maker brings a new breather to the interoperability environment.

When organisations have to be on the market, they have to be ready to satisfy the demands of the customers, satisfy their needs and know what they want, this brings some interoperability issues. In this case a semantic issue, related with the nomenclature and conceptual definition of the chosen products. Occasionally, the customers and the suppliers have their own nomenclatures and their systems should be prepared to align product data and knowledge. To do this, systems must have a dynamic evolution to increase the interoperability level over time, and do an upgrade and enrichment of the reference ontology.

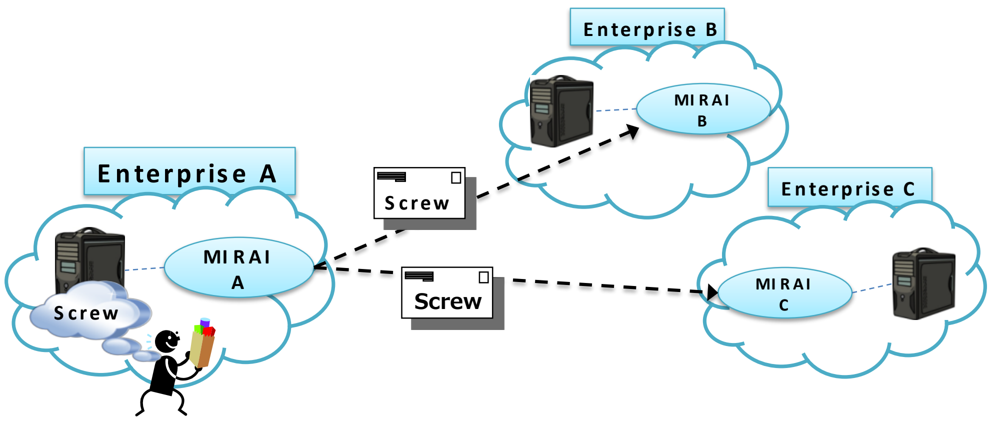


Figure 6.20: Scenario 1: concept not in the reference ontology

The case study here presented, describes a simplified collaborative network between three enterprises that sell bolts. In Figure 6.20, there is a scenario where a client wants to buy a "screw" from enterprise A, however that concept doesn't exist in the reference ontology. The system doesn’t ignore the concept, it records it in the CM as a new searching concept. For each search with this concept its relevance increases. If the system never relates it to other concepts in the reference ontology, and if users still trying to search it, in such a way that it reaches a predefined threshold for this cases, such concept become considered by MIRAI as a “new possible product” on the domain. Thus, MIRAI communicates it to the other network enterprises in order them could readjust their own marketing offers.

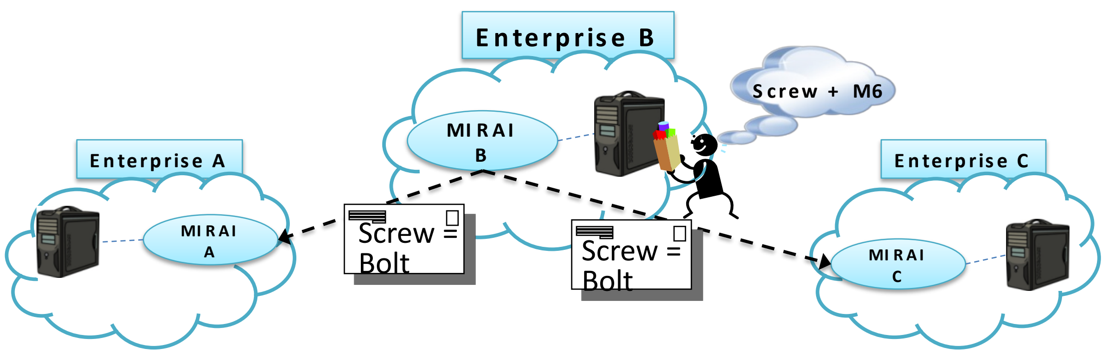


Figure 6.21: Scenario 2: creation of concepts relations

Another scenario is presented in Figure 6.21 where a new customer asks for a “Screw + M6”. In B, the screw is still an unmapped term, but the M6 is familiar type of bolts, so a list of M6 will appear. At this point oLEARCH will associate that a bolt is a screw and will create a relation between them. MIRAI identifies the new morphism connecting bolt with screw, and sends it to the network, to be informed about this morphism.

The last scenario presented in Figure 6.22 illustrates a typical MIRAI sustainability case where the administrator changes the definition of the bolt. He says that a bolt always need to come associated with a nut, and inserts a new definition. So, the MIRAI Intelligent Supervisor block will find this new evolution and proposes a readjustment of the previous mapping (screw = bolt). This is done by creating a new mapping between the new definition of bolt and the screw, with a description saying that a bolt is a specific screw (bolt ⊆ screw). Then it communicates with the other enterprises warning them about this new mapping.

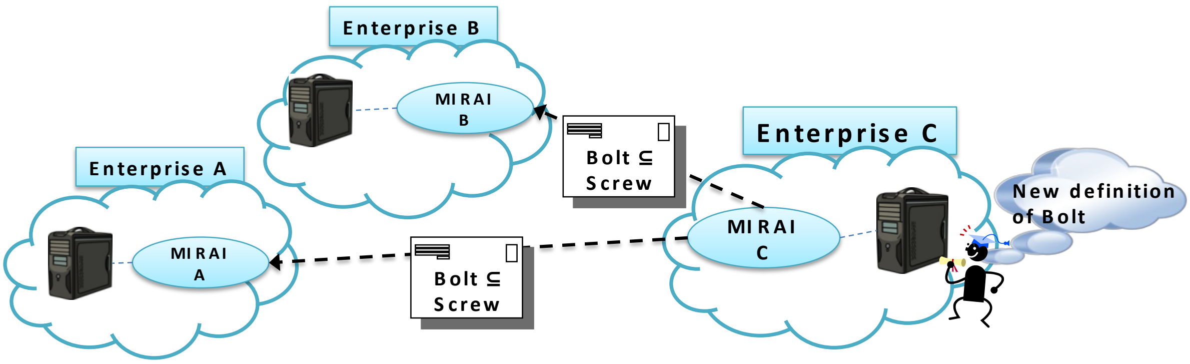


Figure 6.22: Scenario 3: new definition

These three scenarios show the advantages and extra functionalities that the oLEARCH brings to the MIRAI Framework. With the learning ontology implemented (oLEARCH) it is possible to avoid some human intervention in the decision of the mappings, gaining time and money. However, decision makers can still intervene and cause harmonization breaking that needs to be solved with MIRAI.

6.4 Dynamic knowledge Maintenance Plug-in to MENTOR

With oLEARCH the author accomplishes a component with functionalities to manage the dynamic inclusion or exclusion concepts able to enable or assist in the update of a KB. Thus, in the following it is presented an overview of the MENTOR tool architecture accomplished with a dynamic knowledge maintenance plug-in: the oLearch. The proposed architecture (not yet developed) is composed by three main components (Figure 6.23): 1) information models able to represent knowledge; 2) java libraries acting as ontology handlers, machine learning functions, and web services; and 3) user interfaces able to provide MENTOR and its maintenance plug-in functions. The Protégé tools support all of these components namely: Jena API; Protégé serve and its MetaProject Ontology plus the Collaborative Protégé plugin.

With this architecture users through the MENTOR part, are able to create knowledge-based projects or connect to existing ones, and run the entire six MENTOR’s collaborative ontology building steps to the definition of glossary; thesaurus and reference ontology respectively of a specific domain. Other operations have been also developed to provide management operations to the process. The

Metaproject is a frame-based ontology that comes with Protégé. Its key role is to represent the information about hosted knowledge projects with their users and access permissions [24]. The original Metaproject was accomplished with specific MENTOR requirements to keep track of the MENTOR steps and their status, like for instance, to know the name of users that finished the operations of a current step.

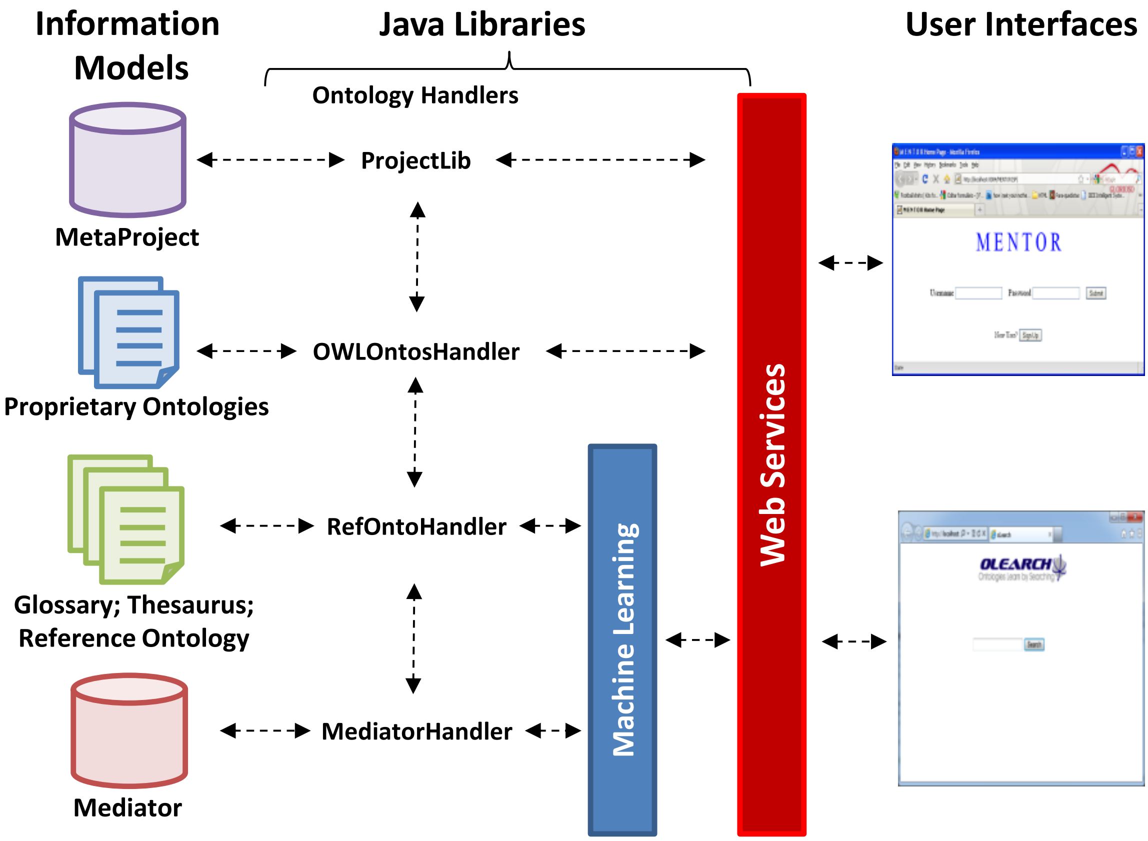


Figure 6.23 – MENTOR with oLEARCH as a plug-in to the Dynamic Knowledge Maintenance

oLEARCH as said before is responsible to provide knowledge maintenance abilities to the proposed architecture. It uses a machine learning technique, which main purpose is to find and describe structural patterns in data [197]. This could be made in support of a product search tool as presented in the use cases presented before. When users search a specific product they introduce concepts that will provide an increase of the lexicon associated to that specific product. Such process is reached through searching, using introduced concepts over the reference ontology to find the desired products classified on that ontology. The machine learning process starts by clustering the introduced concepts with the reference ones (from reference ontology) and, how much times a concept is used, more it gains importance (weight) on decision for the output results. The output is ruled by an instance-based learning approach, which distance is inversely related to the “weight” that the concepts have. This means that the result product is found through a distance function that determines which product is closest in (semantics) relation to the concepts introduced.

After several utilizations patterns are obtained. One example is related to the proposition of new concepts to a specific object. Thus, with proper patterns knowledge can be adapted or maintained to new nomenclatures used by the community. This will be able to be accomplished by a right connection between oLEARCH and MENTOR.

As a final remark, Web Services are useful in this architecture as they are responsible for giving worldwide users the opportunity to use these services, facilitating them the building of their own user interfaces to use MENTOR and oLEARCH functionalities.

This solution proposes a knowledge maintenance stage to MENTOR to improve the management of the knowledge along product life cycle in order to be adapted accordingly to the knowledge evolution of the system. Knowledge maintenance is ruled by the analysis of the users' interactions feedback, which works as the main trigger to the learning process on which such knowledge-based system is based on.

6.5 Conclusions

OLEARCH tool makes use of ontologies with some statistics associated to its concept and operational research methods. OLEARCH also establishes mapping between reference ontology and Mediator ontology concepts. This facilitates to acquire knowledge from users in the sense that their introduced concepts received a statistic number related to its use frequency.

A mapping between a "reference" concept and a "new" concept introduced by a user could be used to propose new reference concepts and products. This represents a knowledge evolution step, where ontology mapping confirmed to be a suitable solution for ontology enhancement. Thus, it could support the implementation of the knowledge maintenance stage to MENTOR. The presented methodology might be applied in any enterprise no matter its domain, as oLEARCH and MENTOR is able to interact between MEDIATOR and any reference ontology.

Since enterprises environments and company's alliances are changing fast in today's competitive market. The knowledge maintenance enabled by oLEARCH assumes a key role to dynamically update reference ontology, and consequently to information systems semantics adaptability.

7 Customisable e-Training

Nowadays with the globalisation phenomenon, companies are pushed to improve its strategies towards deconstruction and a focus on core competencies, giving rise to the concept of distributed virtual enterprises [322]. To reach such competencies, personal knowledge and other intellectual capital assets serve vital functions within the enterprise [323]. Thus, there is an increase demand of workers flexibility and consequently a constant need of delivering training to them [324].

As the economy turns more to a knowledge-based industry, keeping all the members of an organization up to date is becoming an increasing challenge. With the rise of e-Learning, the continuous training is now possible being a major driver to promote the development of competencies and knowledge in organizations.

Distributed virtual enterprises are alliances of organisations that come together to share skills or core competencies and resources in order to better respond to business opportunities [325]. In accordance with this Hamburg et al. stated in [326] that a possible solution is to involve SME's into sharing knowledge and collaboration by building communities of practice and to develop business-oriented models of training to meet their needs.

Human resources and training for achieving business competence, particularly e-competence, represents factors on which competitive advantages are going to be built. E-Learning can further contribute to the achievement of such competences and at the same time can meet the pronounced needs for flexibility in SME's.

Since research carried out in several European and national projects reveals that e-Learning isn't that much applied in small and medium sized companies (SME's) [327]. E-Learning is still used mainly in big companies, where there is sufficient knowledge and resources to develop and implement sustainable training strategies. Only the IT sector and a few more others use regularly e-Learning, but unfortunately the few activities undertaken by SME's usually are of rather poor quality [326], and consequently the effect of the training doesn't often reach the desired outcome.

One possible solution is to involve SME's into sharing knowledge and collaboration by building communities of practice and to develop business-oriented models of training to meet their needs [237]. Also, by delivering content in small pieces over time as part of a large process, corresponding to the needs of SME staff for faster learning in the context of their work, it would be interesting to, by creating a simple application based on a training development methodology, serve several SME's in the training of non private knowledge. Thus, it is important to design effective models and lines of intervention to help SME's to build participative suitable models of training i.e. within communities of practice [237].

Despite setbacks in the past, mostly caused by exaggerated expectations and inappropriate approaches and products, e-Learning is still considered as a key for the solution of the human resources. Consequently, to work towards sustainable learning strategies, e-Learning has to be embedded in intelligent and adequate "mixtures" of different learning methods and technologies [328].

In education, widespread appliance of such a shared instructional vocabulary offers advantages for teachers and learners. A more accurate search for learning resources, made possible by the explicit instructional function, leads to better reuse and less duplication, hence faster authoring of curriculums. By seeking instructionally appropriated learning material, learners can bridge knowledge gaps more efficiently [329].

In order to address this issue several authors suggest that Ontologies can be used to describe LO's content, thus providing LO's with a new dimension of reusability – content reusability. Ontologies for e-Learning are different in the following aspects: content – what the learning material is about, context – in which form a topic is presented, and structure – as learning material does not appear in isolation [330].

Valuable information can be gained by mining metadata of educational resources. However, if the mined data is annotated using IEEE LOM, then significant pedagogical information is missing. While LOM and SCORM provide a framework for the representation and processing of the metadata, they fall short in including the needed semantic density for more specific pedagogical tracking. Using a pedagogical ontology will provide a higher level of decision support analysis and mining, based in qualitative issues like: the pedagogic methodologies used the collaborative degree of activities or the understanding expressed in the assessments [331].

The development and use of conceptual models in learning has been a research topic of the learning sciences for many years. While the earlier works focused on the individual learner, the collaborative use of conceptual models has become a research field in its own later on. Despite the on-going interest in the use of conceptual models for learning, there is a lack of theoretical as well as empirical work regarding the role of ontologies in collaborative learning and knowledge creation [332].

The issues identified before are mainly related to the lack of modularisation, sustainability, adequacy and reusability of eLearning materials or approaches. Author intends to contribute to reach a customisable e-Training programmes approach that could be aligned to these characteristics enabling also the efficiency to reach SME's knowledge assets. Thus, the intention is to have a training strategy aligned with an efficient capability for modelling through ontologies that could support a common methodology definition of building e-Training courses and a semantic enrichment of e-Learning standards.

7.1 E-Training Course's Development Methodology

Closely linked with the way, how a course is developed is the technology that supports it, and knowing that it isn't always needed a complex platform for delivering training [325]. There are several technologies and methodologies that allow the creation of courses and training materials for the industry, but they are mostly academic-focused solutions where the process of building courses and materials it's more exhaustive then what the industry typically requires. This process it's bound to be subject of inefficiencies brought by the using of academic-focused technology in industrial training.

The Training Development translates design specifications into training materials. It starts by identifying the objectives and the target audience including desired roles & competences. Then it uses an appropriate instructional approach to perform the training courses definition and its training materials development, complemented with a set of different quality reviews.

The training development methodology proposed by the author has been validated as a potential solution for improving the efficiency in developing e-Training courses within partners in European research projects as in CoSpaces and CRESCENDO, where the exchange of information results in a need of constant training between the several parts involved, by structuring the training developing procedure. The training development methodology follows a specific process, composed by three different task tracks (training development, overall training validation and training execution) that complement each other, to reach the training implementation see Figure 7.1.

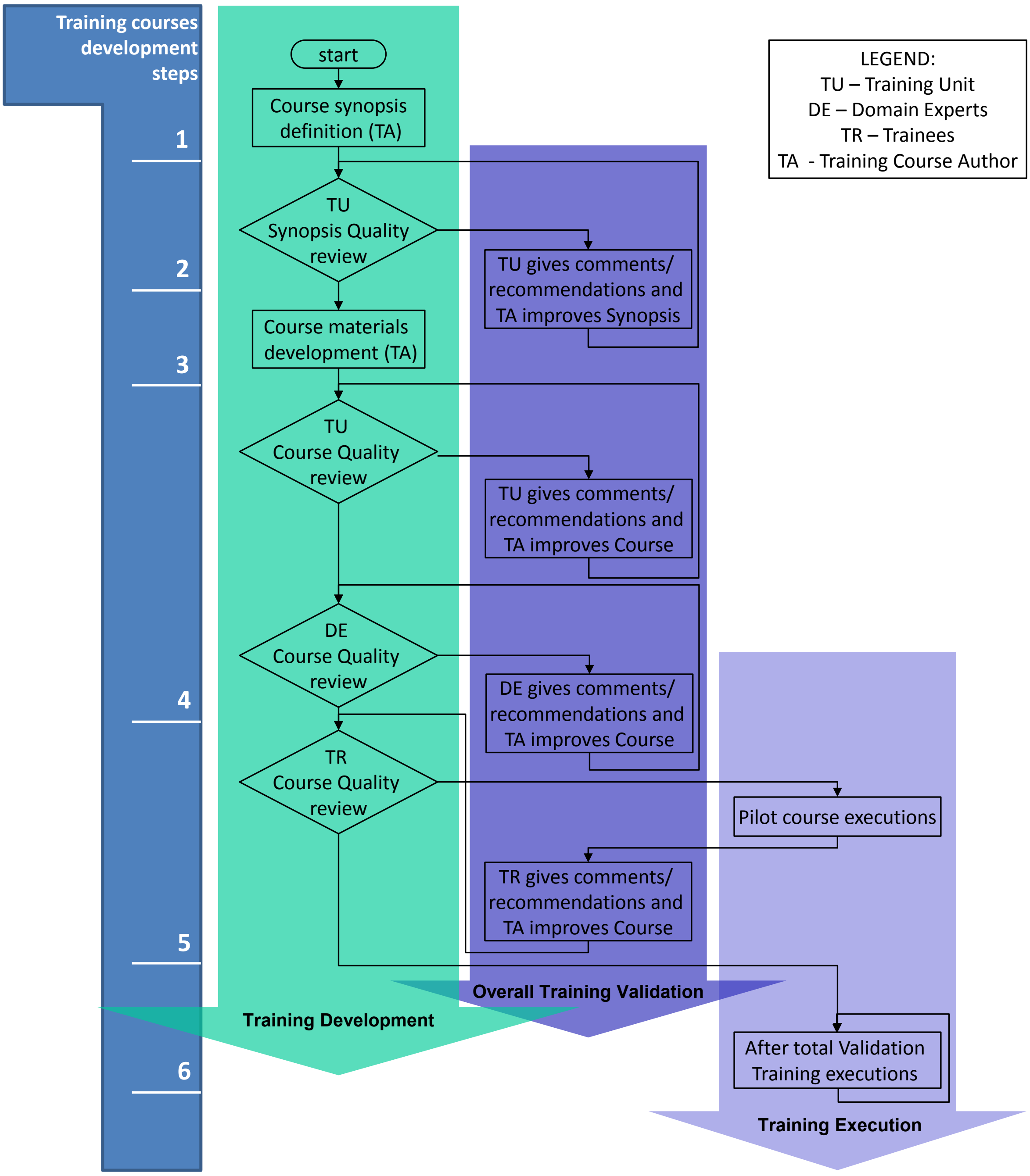


Figure 7.1 - Detailed Training Implementation Process

The training development starts by defining the course's synopsis according to the training overall objective. Then the Training Unit (TU) performs a quality review of the synopsis. This is done in the training validation track. Thereafter, the course material is developed, followed by two more quality review cycles, similar to the first. The first is made internally in the TU. Training Experts (TE) who are familiar with the training contents and able to analyse and validate the courses perform the second cycle. The Training Authors (TA) improves the courses based on the feedback from these two validation activities. After this, in the next phase, another quality review is performed through a pilot course execution. Here trainees are the ones who give comments about the courses and the TA once again improve the courses for the final release.

The courses should be at this point, able to train a specific audience, about some specific topics. This should reach the purpose to deliver trainees the appropriate skills able to help them to reach specific identified competences. All of these facts should be aligned to the courses synopsis defined at the beginning of this entire process. All this process is accomplished by following a determined steps status shown at the left part of the Figure 7.1 and detailed at Table 7.1.

Step	Activity	Artefacts available / used on the activity
1	training course/topic identified	A line in the list of courses
2	course owner accepted training course proposal	Synopsis + Early raw material defined
3	Defined contents and agreed course structure	Early material organised according to the agreed couse structure – draft version of the course
4	Training course in development	1st full version of the course
5	Training execution	Feedback forms
6	Improve course based on feedback	Course updated - final version of the course

Table 7.1 – Steps of course development

7.1.1 Course Synopsis definition

The training development intends to be aligned with the general training objectives. The training policy should be aligned with participants companies training policies, answering to the skills and competences required for a specific overall objective as for instance, interoperability solutions implementation. Consequently, it is needed to develop, organise and run training courses, to train the "future users" (in a broad sense) to interoperability solutions use, implementation and support.

The process starts by defining the training target audience profiles or types and required roles & competences. Then a procedure—composed by 5 steps—to assist with the definition of the courses synopses:

Step 1 – Training Topics Definition

This step helps people identify topics that should be addressed in the training courses. The topics are defined by answering the following questions:

- **Question 1.1** - What training topics (discussed under this integrated tool / team / group) should/could be used for training purposes?
- **Question 1.2** - What do you envisage will be the new knowledge (or know-how) from

application / deployment of results from this tool / team / group?

- **Question 1.3** – In your own words, which will be the new/'updated' competencies from this work?

Step 2 – Target Audience and Training Course Types Definition

This step is related to the identification of the target audience and competences, to which the training topics are related. Thus, the questions to answer in this step are:

- **Question 2.1** - What do you envisage is the target audience for each of the training topics identified in step 1?

- **Question 2.2** - What do you envisage to be the associated competences and/or roles that are enabled by those training topics?

Step 3 – Training Objectives Definition

This step is related to the training objectives definition for each training topic identified. The training objectives are represented by the soft and hard skills that a trainee should learn from the training topic that will help him/her reach the related competence.

- **Question 3.1** –> For a training topic without associated competences: What should trainees know, understand, and be able to do on this training topic?

- **Question 3.2** –> For a training topic with associated competence(s): What should trainees know, understand, and be able to do on this topic in order to reach the identified competence?

Step 4 – Training Themes and Structure Definition

Two sub-steps of tasks compose step 4:

- **Sub-step 4.1** -> Organise a list of groups of {topic ideas; statements; competences; and training objectives} related to each other (identified in the steps before). Each of these groups is at the first instance, a rough, and potential, training module.

- **Sub-step 4.2** -> Try to identify a more generic theme that could categorise and consequently aggregate one or more of the rough training modules identified. Note that each theme identified could represent a potential course. Thus a theme that aggregates all the rough training modules could be the most appropriate.

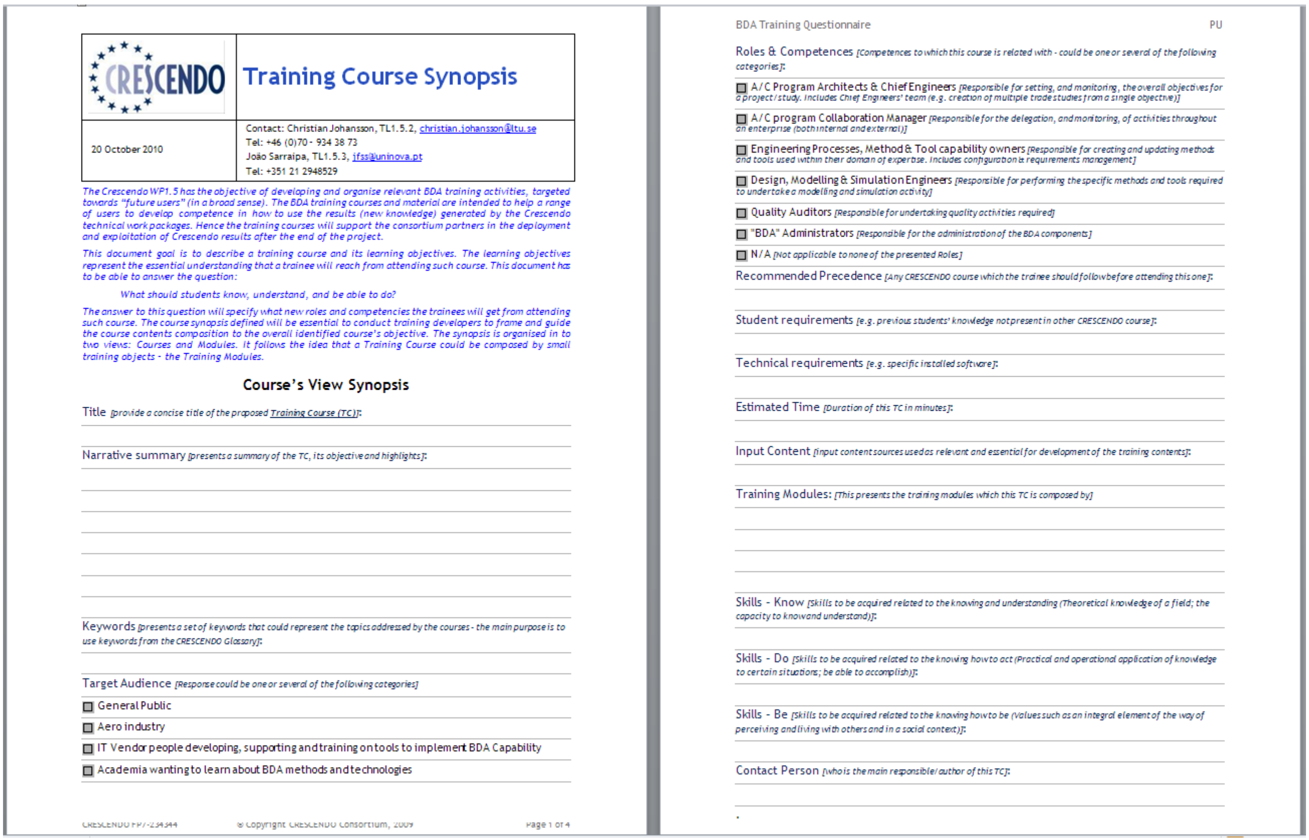
- **Sub-step 4.3** -> For each rough training module, categorised around a specific theme, the objective is now to organise them to establish an order of execution. This is a way to establish a course with an organised set of training modules around the theme. Thus, each rough training module could become a final training module, constituting in such way a training element of the training course structure. A course structure should typically be initiated by a "Module 0" that is the introduction to the topic. "Module 1" and following ones (except for the last) are related to the content itself. Finally the last module is the conclusions module (see

Figure 2.25), which summarises the main points and conclusions from the course.

Step 5 – Training Courses Synopsis Definition

For a training author it is easier to produce small pieces of training elements in order to be more organized. First the author starts with a global training objective (course objective), which then is divided into smaller objectives (training modules objective). With this, the reference training courses contents are divided into manageable modules to facilitate the organization in producing courses. Moreover, such training modules can be reused separately one another in an atomic way by other trainers (e.g. academics in their classes as additional educational materials).

This step defines the course using the previous identified resources/ideas by completing a "course synopsis" (Figure 7.2 - A Training Course Synopsis Example), which is composed by several sections; course title, narrative summary, recommended precedence, estimated time, target audience – which relates to an intended trainee profiles, and etcetera. Annex 10.2 features an example excerpt of a course synopsis. It presents explanations and clarifications for each synopsis section (e.g. Skills – Know [Skills to be acquired related to the knowing and understanding; Theoretical knowledge of a field; the capacity to know and understand]).



CRESCENDO	**Training Course Synopsis**
20 October 2010	Contact: Christian Johansson, TL1.5.2, christian.johansson@ltu.se Tel: +46 (0)70 - 934 38 73 João Sarraipa, TL1.5.3, jfss@uninova.pt Tel: +351 21 2948529

The Crescendo WP1.5 has the objective of developing and organise relevant BDA training activities, targeted towards "future users" (in a broad sense). The BDA training courses and material are intended to help a range of users to develop competence in how to use the results (new knowledge) generated by the Crescendo technical work packages. Hence the training courses will support the consortium partners in the deployment and exploitation of Crescendo results after the end of the project.

This document goal is to describe a training course and its learning objectives. The learning objectives represent the essential understanding that a trainee will reach from attending such course. This document has to be able to answer the question:

What should students know, understand, and be able to do?

The answer to this question will specify what new roles and competencies the trainees will get from attending such course. The course synopsis defined will be essential to conduct training developers to frame and guide the course contents composition to the overall identified course's objective. The synopsis is organised in to two views: Courses and Modules. It follows the idea that a Training Course could be composed by small training objects - the Training Modules.

Course's View Synopsis

Title *[provide a concise title of the proposed Training Course (TC)]*:

Narrative summary *[presents a summary of the TC, its objective and highlights]*:

Keywords *[presents a set of keywords that could represent the topics addressed by the courses - the main purpose is to use keywords from the CRESCENDO Glossary]*:

Target Audience *[Response could be one or several of the following categories]*

- ☐ General Public
- ☐ Aero industry
- ☐ IT Vendor people developing, supporting and training on tools to implement BDA Capability
- ☐ Academia wanting to learn about BDA methods and technologies

CRESCENDO FP7-234344 © Copyright CRESCENDO Consortium, 2009 Page 1 of 4

BDA Training Questionnaire PU

Roles & Competences *[Competences to which this course is related with - could be one or several of the following categories]*:

- ☐ A/C Program Architects & Chief Engineers *[Responsible for setting, and monitoring, the overall objectives for a project/study. Includes Chief Engineers' team (e.g. creation of multiple trade studies from a single objective)]*
- ☐ A/C program Collaboration Manager *[Responsible for the delegation, and monitoring, of activities throughout an enterprise (both internal and external)]*
- ☐ Engineering Processes, Method & Tool capability owners *[Responsible for creating and updating methods and tools used within their domain of expertise. Includes configuration & requirements management]*
- ☐ Design, Modelling & Simulation Engineers *[Responsible for performing the specific methods and tools required to undertake a modelling and simulation activity]*
- ☐ Quality Auditors *[Responsible for undertaking quality activities required]*
- ☐ "BDA" Administrators *[Responsible for the administration of the BDA components]*
- ☐ N/A *[Not applicable to none of the presented Roles]*

Recommended Precedence *[Any CRESCENDO course which the trainee should follow before attending this one]*:

Student requirements *[e.g. previous students' knowledge not present in other CRESCENDO course]*:

Technical requirements *[e.g. specific installed software]*:

Estimated Time *[Duration of this TC in minutes]*:

Input Content *[input content sources used as relevant and essential for development of the training contents]*:

Training Modules: *[This presents the training modules which this TC is composed by]*

Skills - Know *[Skills to be acquired related to the knowing and understanding (Theoretical knowledge of a field; the capacity to know and understand)]*:

Skills - Do *[Skills to be acquired related to the knowing how to act (Practical and operational application of knowledge to certain situations; be able to accomplish)]*:

Skills - Be *[Skills to be acquired related to the knowing how to be (Values such as an integral element of the way of perceiving and living with others and in a social context)]*:

Contact Person *[who is the main responsible/author of this TC]*:

Figure 7.2 - A Training Course Synopsis Example

7.1.2 Course Materials Development

Using the training course synopsis, authors develop the course modules contents using materials developed for the interoperability solutions description and, when needed, complemented using external materials. Training materials development is related to all the actions needed to produce materials to training implementation. A structure with assessments included is a good approach for producing an interactive course. Thus each module should have at least one assessment question to trainees test their self-progress in learning. In addition, each course should have questionnaires to obtain feedback first from training experts and then from trainees to accomplish the course validation and evaluation and thus improve the quality of the materials.

Training the authors

Authors receive a special course to train themselves authoring and developing their courses. Such authoring sessions are like seminars where the authors receive some directives and explanations in how they should prepare their training materials in order to have all the materials developed using the same structure and approach.

7.1.3 Overall Training Validation

The **training validation and evaluation** ensure that training-under-development stays on track, safeguarding achievement of training goals and analysing system performance. A quality review process based on decisions and revisions for future course iterations can be made after evaluating the strengths and weaknesses in a completed training programme, thus ensuring achievement of desired goals.

Training Validation

Training Validation certifies that courses follow its own synopsis and accordingly to general quality aspects defined for the purpose. The quality aspects to be assessed by the training courses validation and evaluation include:

1. **Correctness / Reliability / Clearness** - To ensure that all training material is correct and reliable according to the training objectives and high quality standards.

2. **Efficiency / Suitability** - This term was taken to ensure the suitability of each of the training package towards the related target audience. To have appropriate trainees' self-assessment to improve their self-learning effectiveness. Adding to this, regarding post-update of the training material, suitable questionnaires to be used to gather trainees' feedback to subsequent improvement.

3. **Modularity / Flexibility / Portability** - All training material has to be developed in modules so that they can be re-used or integrated in various training units and for various purposes (training programmes).

4. **Coding Standards** - Proper coding standards followed (e.g. synopsis definition) for the training materials development to ensure an easily comprehensible and highly effectiveness of the training executions.

5. **Functionality / Robustness** - The training material functionality should be operational under various conditions. The material should be developed in the most common accepted software formats.

6. **Intellectual Property Rights (IPR)** - All material developed has to be carefully reviewed to ensure that any material released to the public is audited.

The training validation is the verification process of the training development. It is accomplished through two quality review cycles: 1) the training unit validation, and 2) the training experts' validation. This process is called validation because it validates the courses, if they are or not ready to proceed to the next training phase, the training execution.

The **training unit validation** verifies the courses ensuring that they have its materials compliant with all the quality aspects presented before except for IPR, which is covered by the TE. Since training unit members have a clear idea about all the courses and their role on the overall training organisation, they are able to check their contents in such way to ensure their general quality and purpose.

The **training experts'** training validation ensures that courses follow mainly three of the quality aspects defined. These aspects are related to: 1) the correctness, reliability and clearness of the contents; 2) the suitability of each of the training package towards the related target audience plus competences; and 3) the IPR issues.

The training experts are the ones that have a wide knowledge about what should be present in courses to help a specific target audience develop a predefined set of competences. Thus they can properly validate the course contents suitability and their clearness. In additional, they are able to analyse the possible IPR issues that courses could have.

Training Evaluation

The training evaluation definition is concerned to the trainees' act of formulating a judgement that in this case, it is on the training course execution. This definition it was defined based on the Hurteau et al. in [333], which states that the main purpose of a program evaluation can be to "determine the quality of a program by formulating a judgment".

Therefore, in each training execution the trainees are asked to feedback concerning, e.g. materials quality, trainers performance, etc. This information is used for a constant judgement of the courses and their subsequent improvement.

The feedback mechanisms have been implemented through feedback questionnaires to be filled by trainees at the end of a training activity (training course). These questionnaires can be composed for instance by a set of questions as the following ones: 1) Were the objectives achieved?; 2) Were the concepts and techniques clearly explained?; 3) Was the course structured in a logical way? ; 4) Was the course easy to follow?; 5) Were you encouraged to actively participate during the course?; 6) Did you like the lesson time frame?; 7) What else do you think should be included/excluded in the course?; 8) Were the topics presented relevant to your work?; 9) Was the course interesting and enjoyable?; 10) Would you recommend the course to others?; 11) Do you believe the skills you have learned will help you improve your performance on your job?; 12) Are you interested in using the course materials as a reference?; 13) Do you think the course should be good value for money?; 14) Were your expectations met?

Trainees are invited to answer the above questions using one of the following five possible statements: - Not at all; - Maybe Not; - Not Sure; - Maybe Yes; - Definitely yes.

There are some questions that invite trainees to state some considerations about the course as the number seven of the above list.

7.1.4 **Training Execution**

The training execution is the act of delivering training about a specific theme/domain with a comprehensive purpose to enrich the knowledge and/or to provide precise competences to a pre-determined target audience.

The training execution phase focuses on setting up the training environment, and delivery or distribution of instructional materials ensuring delivery of a training session able to capture trainees' interest.

Normally training executions are performed using a blended learning approach. Therefore, it accomplishes a full set of solutions: traditional classroom, virtual classroom and web-based training supported by e-Learning material for interested parties within the identified profiles. Recognised trainers supported by specific skilled experts should conduct the training sessions. This could facilitate the training execution through different training types as:

1. **Bound by time and place** - such as classical classroom and instructor-led training; is the traditional training delivery method that requires the physical presence of students at the location at which the training takes place. This type of training is particularly suitable for contents that impart a substantial understanding of the subject and require intensive interaction with the instructor concerning further inquiry, practical exercises, or collaboration between course participants.

2. **Bound by time only** - e.g. instructor-led synchronous virtual classroom sessions (webinars), where a tutor guides the learning activities over Internet; is an alternative training delivery way that has to be supported by a specific platform. It provides many of the advantages of classroom training without the constraint of a fixed location.

3. **Both independent of time and location** - such as web-based courseware that the learner can access ubiquitously, i.e. anytime and anywhere. This training delivery type allows course participants to extend the training time-span according to their needs and within a specific personal time al-location. The learning objectives achieved can be validated by built-in assessments that allow electronic marks through the evaluation of multiple choice questions. Web-based training offers the highest flexibility but also requires an additional discipline because is the user that controls the organization of the learning process.

7.1.5 **Training Implementation**

The Training Implementation is in a broad sense related to all the activities needed to perform training delivery or execution since the beginning, until the last step. Thus, its analysis is related to the work conducted since the training courses definition, until the delivery or distribution of such instructional materials ensuring training sessions able to captures trainees' interest.

7.2 E-training Implementation Case Study and Demonstration

The case study is related to a training implementation need of the CoSpaces research project. CoSpaces is an IP project funded by the EC under the IST Programme of the FP6, which overall objective is to develop organisational models and distributed technologies supporting innovative collaborative workspaces for individuals and project teams within distributed virtual manufacturing enterprises [233]. Training in CoSpaces project aims exactly at providing knowledge and skills that allow key target groups within distributed manufacturing enterprises to understand collaborative practices and acquire the practical experience of collaborative design engineering methods, supported on meaningful case studies and demonstrators [230].

The training curriculum, in this case study, was designed to establish the underpinning that is to be used to frame down training courses and, subsequently, the automatic orchestration of courses. The Training in Collaborative Working (TiCW) follows an integrative structure established over two dimensional axes: horizontally, training levels; and vertically, reference content areas. The Figure 7.3 illustrates how the TiCW Curriculum as the training levels classified in three stages [230]:

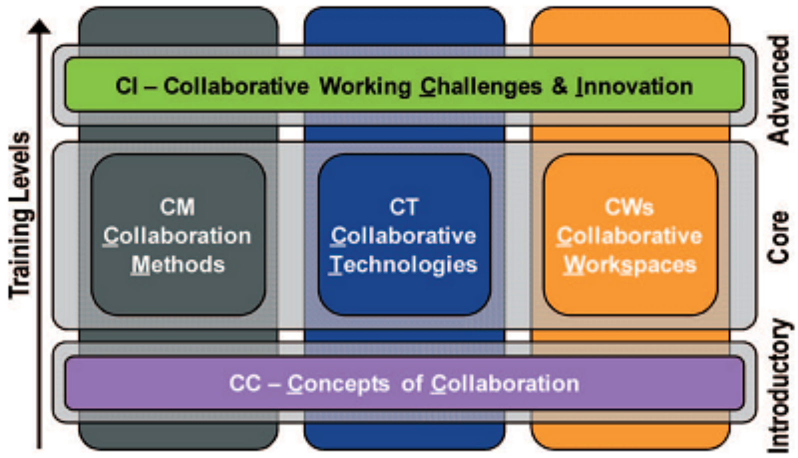


Figure 7.3 - TiCW Curriculum

1) **Introductory** level, with the objective to endow trainees with the basis about the curriculum areas and to ensure a common knowledge level amongst all;

2) **Core** level, which aggregates training elements that establish the in-depth knowledge, understanding and skills that are central to the curriculum areas;

3) **Advanced** level, combining training elements focusing on advanced topics that provide enhanced comprehension of specific knowledge areas related to collaborative knowledge areas related to collaborative working pragmatics.

The curriculum content areas have been defined upon reference subjects that represent the fundamental dimensions of collaborative working pragmatics: 'Collaboration Methods', 'Collaborative Technologies' and 'Collaborative Workspaces'. In conjunction to these areas there are two more that are related to the introductory and advanced levels of the curriculum which are designated by ‘Concepts of Collaboration’ and ‘Collaborative Working Challenges & Innovation’ areas respectively.

In order to have a better idea of what kind of contents are expected in each course and what is its role in the overall training, the courses were categorized in the curriculum structure as it is shown in the figure bellow.

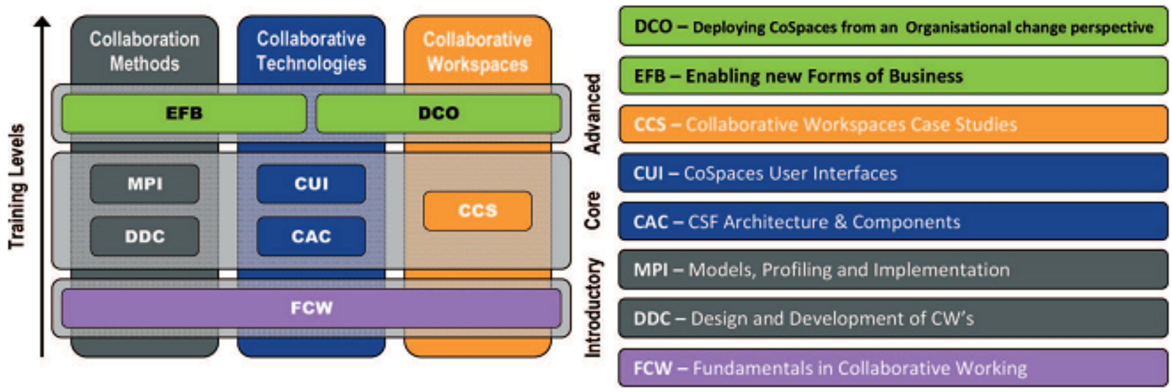


Figure 7.4 - TiCW Curriculum with detailed courses

By establishing a matrix of courses, the curriculum creation allows the prioritizing of courses based on its training level and content area. It is therefore vital for the automatic orchestration of courses, because without a matrix, the correct organization would be a greater challenge.

All the courses are directly related in terms of contents to its curriculum position (training level & content area) and industrial directive. As an example, the FCW - *Fundamentals in Collaborative Working* course belongs to the introductory level of the TiCW curriculum. Thus, its main objective is to endow the students with the basics and scope of Collaborative Working. It is designed as an introduction to any participant in the CoSpaces training curriculum. This specific course has 6 modules as described in Figure 7.5.

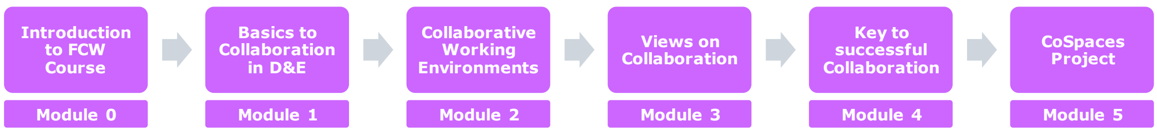


Figure 7.5 - FCW Course Module’s structure

The modules are the atom unit of the course model. Each module has associated keywords to be used in discovering, and reasoning of associated elements, as modules with precedence, courses, and so on. Such information was introduced on the courses synopsis.

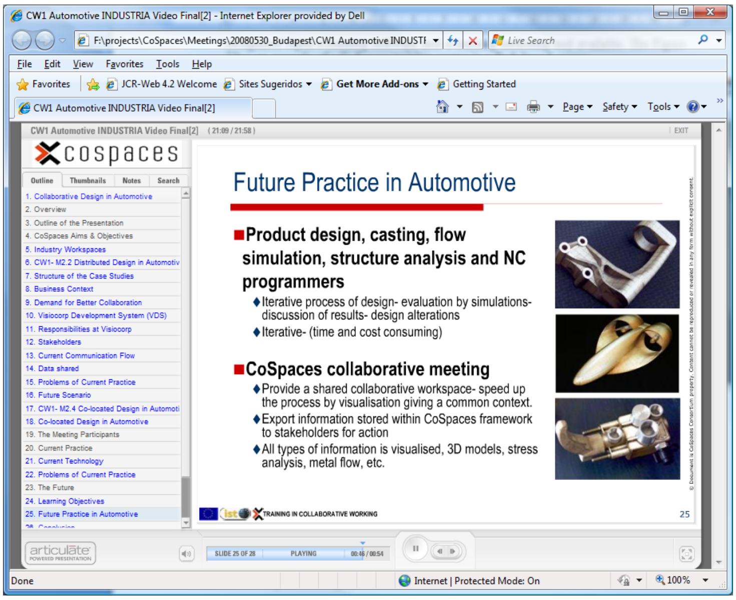


Figure 7.6 - Example of training module in flash

To each module there is a corresponding flash video (Figure 7.6), and subsequently a link to access it. By providing a set of links to a user, he/she could navigate though them and gather knowledge about the selected topics.

7.2.1 E-Training Knowledge Modelling

An ontology can be used to represent a training KB, facilitating the categorization of its elements and subsequently reasoning over it. Below is presented an illustration of a training system. This model was built in Protégé and follows the e-Training course's development methodology defined at sub-chapter 7.1. Figure 7.7 presents an illustration of this use case e-Training ontology, thus, it represents its knowledge model.

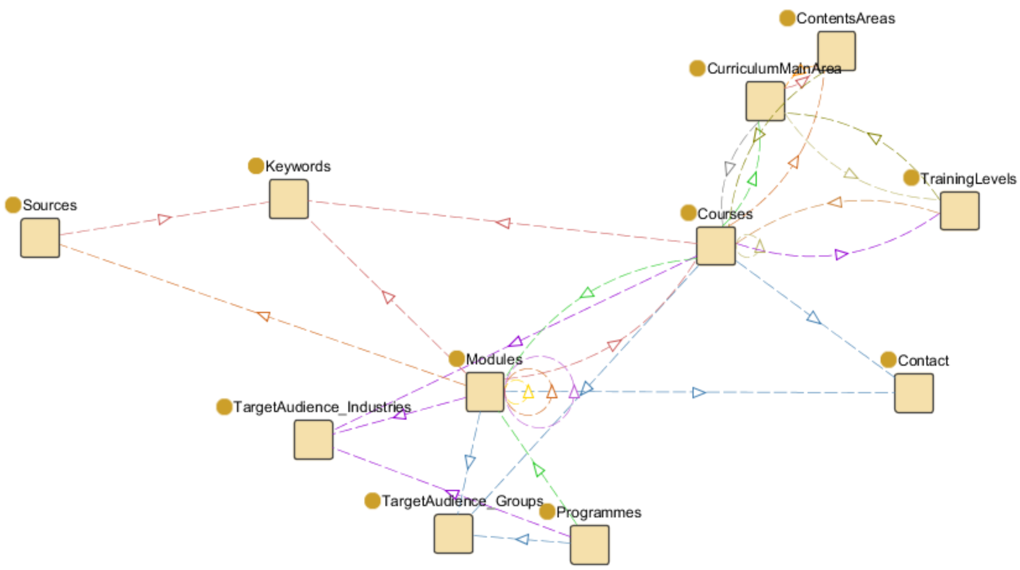


Figure 7.7 - Relationship structure of the e-Training Ontology. Each colour represents a different property linking the concepts. Example: each Module has a set of Keywords, the property linking it is: hasKeywords

In this ontology (Fig. 5) each learning Module has several concepts associated, the Sources concept contains information about the sources referred in the Module, Contact includes the contact information of the author of a Module or Course and Keywords that contain a list of all relevant keywords needed for describing the contents of the Module. A Course, other than Contacts and Modules that contain the course also includes Keywords (that include Keywords inherited from its Modules) and belongs to a Curriculum Main Area that is divided by Content Areas and Learning Levels. Each Module and Course has a Target Audience Group and a Target Audience Industry, to be recommended accordingly to the profile of the learner. Finally, a pre-defined Programme is defined for a specific Target Audience Industry and Target Audience Group.

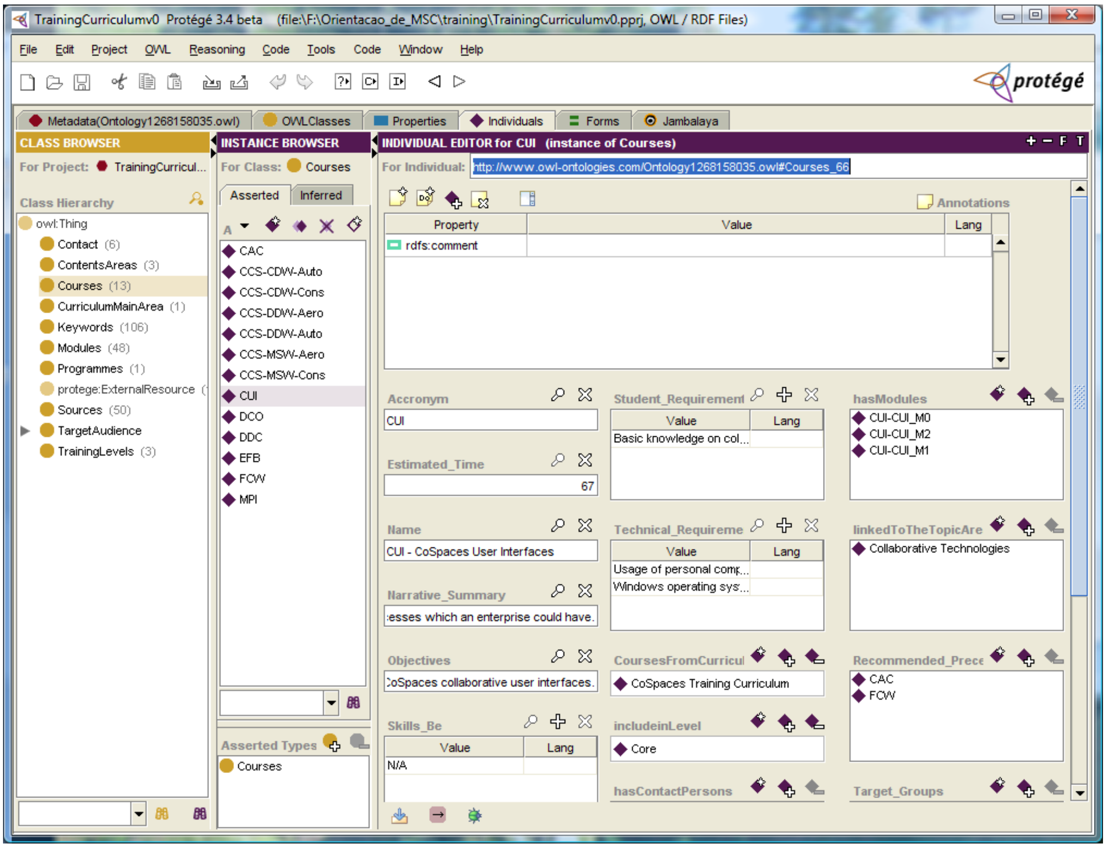


Figure 7.8 – Training System KB in Protégé

The model provides an easy comprehension and management of the whole system by adopting a visual knowledge modelling approach for reasoning and inspection based on Protégé. Figure 7.8 presents the knowledge related to one of the modelled courses – the CUI course. It also shows that the CUI course has 3 modules, and it is linked to the curriculum topic area Collaborative Technologies. The recommended precedence's and the level of training to which the CUI course belongs are also in the model (Figure 7.8), among other relevant concepts.

7.2.2 **Customisable Training Programme Service**

The service generates adaptable courses with content originally developed for static courses. By having a training curriculum matrix and related data represented in an ontology, it is possible to reason over it and generates a training programme, which contents (courses / modules) are presented according to the user needs and following pre-determined pedagogical directives. Figure 7.8 presents an illustration of the e-Learning model that has such training curriculum represented and which was built in Protégé.

The implementation architecture is composed by three different parts: 1) the Training Orchestration Server; 2) the JSP files used for interface; and 3) the server with the flash videos.

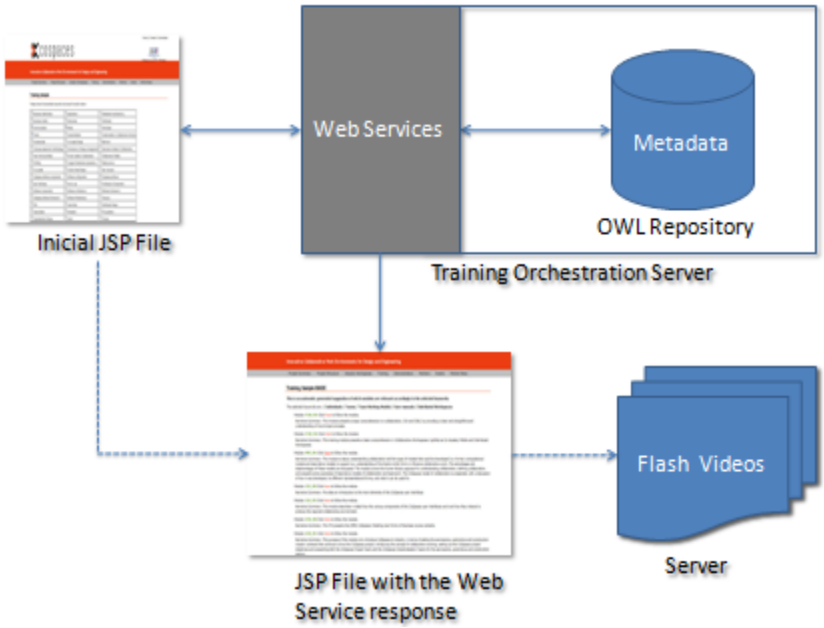


Figure 7.9 – Implementation Architecture

With this architecture a user is able to retrieve an automatic orchestrated course that is built accordingly to its needs. After connecting to the web service of the Training Orchestration Server in order to fetch the available keywords and target audiences, the user selects and sends to the web service the intended fields of search. By reasoning over the Metadata of the Training Systems Modelling existing in the OWL Repository, it is generated a list of modules suggested to be followed and the answer replied and presented in the final JSP file. The response includes all the links to the modules suggested, that is in another server.

The mentioned service is available at http://gris-public.uninova.pt:8080/cospaces/ATPS_in.jsp (website accessed at March 2012). The service is available through a JSP page to the users, presenting a list of all the available training keywords. The user can select the ones that he/she could be interested in. After a submission, a training programme is prompted on the fly, adapted to the choice. A list of training modules is then presented. With such a list, the trainee could follow a training programme adapted to his/her interests. The Figure 7.10 (left part) illustrates the first step: the user selects a set of concepts. In this case, the concepts are: 'Challenges'; 'Collaboration', and 'Collaborative technologies'. After pressing the available button, a suggested training programme is displayed, composed by all the modules that contain the selected keywords and the recommended precedence's of them. On this sample, the trainee should start its learning process by a module that "presents a basic comprehension on collaboration, CW and CWE (…)" as shown in the first training module description of the list that resulted from such process (right part of Figure 7.10).

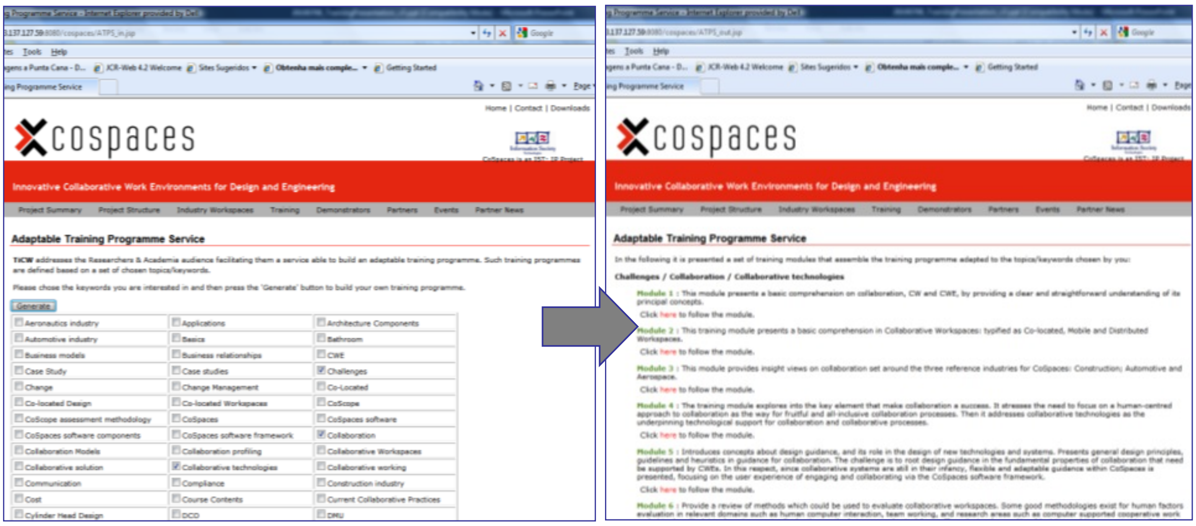


Figure 7.10 – Training orchestration example using the Adaptable Training Service

7.3 Conclusions

E-Training is a major driver to promote the development of competencies and knowledge in enterprises. A lack of customizable e-Training programmes based on trainees' profiles and of continuous maintenance of the training materials prevents the sustainability of industrial training deployment and enterprise knowledge assets enhancement. This chapter enriches the presented training system in sub-chapter 3.3 with a methodology to build training courses. It intends to facilitate the management of all the training components and tasks to be able to build a training course focused in a specific planned objective and enabling various organizations to participate actively on its production. It is aligned with the overall training approach presented, which intends to implement training within the scope of fundamental properties such as dynamism, adaptation and expandability.

Thus, it follows the idea of a dynamic training curriculum, which pursues a modular approach holding atomic training modules that allow the same training contents to be reused and composed in the best manner to reach each training audience's profile or objective.

Since formal KR through ontologies enables computers characterize relevant information related to determined process elements and actors in a domain for specific advanced reasoning. The modelling approach followed enables the creation of supporting methods and tools that will enable the development of dynamic on-demand assembly of modules to promptly develop specifically-targeted courses and training programmes based on the trainees' profiles assessment.

The main message to retain from this chapter is that ontologies, standards specifications and structured methodologies help on knowledge organisation facilitating its systems intelligence increase, enabling the support to the generation of advanced and customisable e-Training services for all. Thus, the knowledge organisation is one of the most important issues to be addressed when a solution for generating instructional designs is introduced.

8 Dissertation Analysis, Assessment and Validation

This chapter objective is to analyse some points or conclusions of the research conducted, checking what semantic related topics were addressed. An assessment of the research approach followed is also conducted to prove the thesis proposed and to evaluate all the dissertation achievements.

8.1 DSA implementations analysis and validation

MENTOR presents a reference ontology building methodology that mainly contributes with functionalities examples to the KA and KMo stages of the SIE abstract system (Figure 8.1). In such ontology building process, it is always addressed KA features when it is asked to participants to contribute with their knowledge, e.g. in concepts and related definitions acquisition. And it is addressed KMo features when the methodology provides guidelines or support to the glossary, thesaurus or ontology building process. It represents an action of modelling the existent knowledge (KMo).

The proposed knowledge architecture for multicultural manufacturing systems (chapter 5) organises knowledge of the furniture domain in various KRE (e.g. dictionaries; ontologies; etc.). Such research work has the main objective of addressing multiculturalism and semantic enrichment of product catalogues, namely AP236 standard. Multiculturalism on the proposed architecture is represented by the acceptance of views (e.g. nomenclatures) in an only domain, and by the multilinguism addressed. The way knowledge is organised by this multicultural approach is obviously related to KMo stage. The semantic enrichment of product catalogues even being related to the KMo topics is mainly related to the use of knowledge by the furniture industry. The semantic enrichment has the main objective of putting in utilisation the knowledge acquired by the industry mainly to the classification of products. Consequently KMo and KU are the main stages of SIE addressed on this knowledge-based architecture (Figure 8.1).

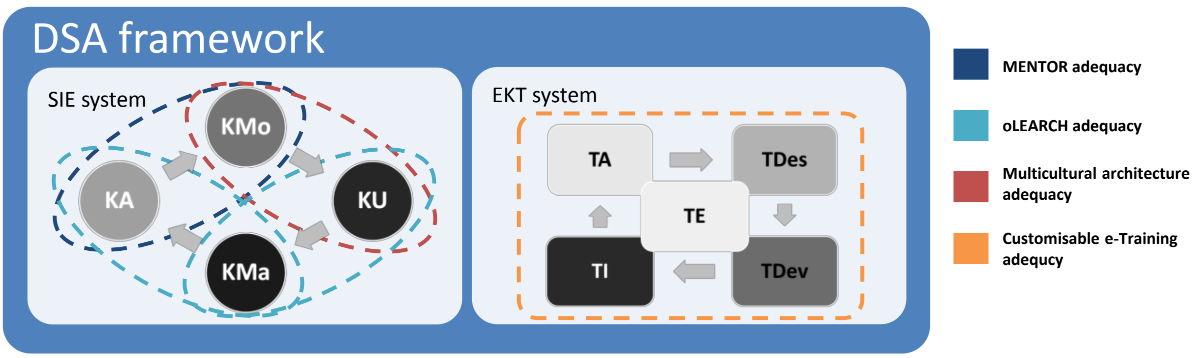


Figure 8.1 – DSA Implementations adequacies

OLEARCH presents a product search engine technically focused in an ontology based paradigm, which states that ontologies learn by searching. This specific engine "uses" knowledge represented in ontologies to answer to products search. However it aggregates the concepts used by the users to learn new conceptual connections updating its weight accordingly to the users searching patterns. Thus, in additional to its "knowledge acquiring" process is able to update its own KB based on such patterns, representing in that way a KMa of its system. As a consequence of these characteristics it is concluded that this prototype addresses mainly the KU, KMa and KA of the SIE stages (Figure 8.1).

Customisable e-Training orchestration proposal is supported by advanced searching or reasoning features resulted from the capability of having an ontology representing the various LO and associated characteristics. Ontologies are able to be used to represent the associated e-Training knowledge due to the efficient training system implementation organised accordingly to the ADDIE ISD and to a structured e-Training course development methodology. Therefore the overall approach for the customisable e-Training orchestration addresses all the EKT stages as illustrated in Figure 8.1.

All the DSA identified main components (SIE and EKT stages) were researched resulting in several approaches methodologies or prototype that could support a specific semantic interoperability solution. However, integration between all developed components still to be accomplished, thus, to have a DSA system able to work as a standalone tool needs further work. Consequently, having DSA as a system able to dynamically adapt to conceptual changes evolution occurred on the network is the desired point to reach, but it should be solved in the near future. To have a DSA able to adapt to any changes with certain wisdom, putting by itself the adaptation in a philosophical probe is another question, and that remains in a distant future.

In the research work conducted to the DSA implementation it was verified the need to technically validate the result of a conducted semantic adaptation result and a business validation of any semantic interoperability resolution approach. Accordingly to this author developed a semantic checking framework and an economic viability methodology.

8.1.1 **Semantic Checking Framework**

Consistency is defined in the Oxford dictionary [334], as the quality of achieving a level of performance, which does not vary greatly in quality. This can be interpreted as something that has an accordance with previously stated facts or characteristics. The consistency of an ontology can be defined as incorporating new information in accordance to the one that was previously represented in the ontology. As ontologies evolve incorporating modifications of the application domain, additional functionalities according to changes are needed. Thus it is important to have a mechanism that can validate that the information within the ontology remains consistent. Much work has been done in this field, such as, frameworks that provide strategies for detecting and repairing inconsistencies [335] using descriptive logic reasoning tools to infer logical consequences, through inference engines, based on a set of rules or facts. Examples of consistency checking tools and inference engines are ConsVISor [336], FaCT++ or HermiT [337]. This type of consistency checking is related to semantic checking, which refers to the validation of ontological structure and concepts semantics.

According to Li et al. in [338], there are three types of semantic checking, single, composite and multiple (Figure 8.2). In the first case the semantic checking is done within a single ontology and it is only deemed consistent if it satisfies a set of concepts and axioms and if all used entities is defined. The second type refers to the semantic checking of ontologies (or subsets of ontologies) within ontologies. Also, in this case an ontology is deemed consistent if the ontology itself and all its included ontologies are consistent. Finally the third type is the main focus of this work and depicts a scenario where several separate ontologies interact with each other. In this case of multiple semantic checking the goal is to validate if all knowledge represented in a given ontology can be represented in another (within the same domain), by means of a reference ontology, for example.

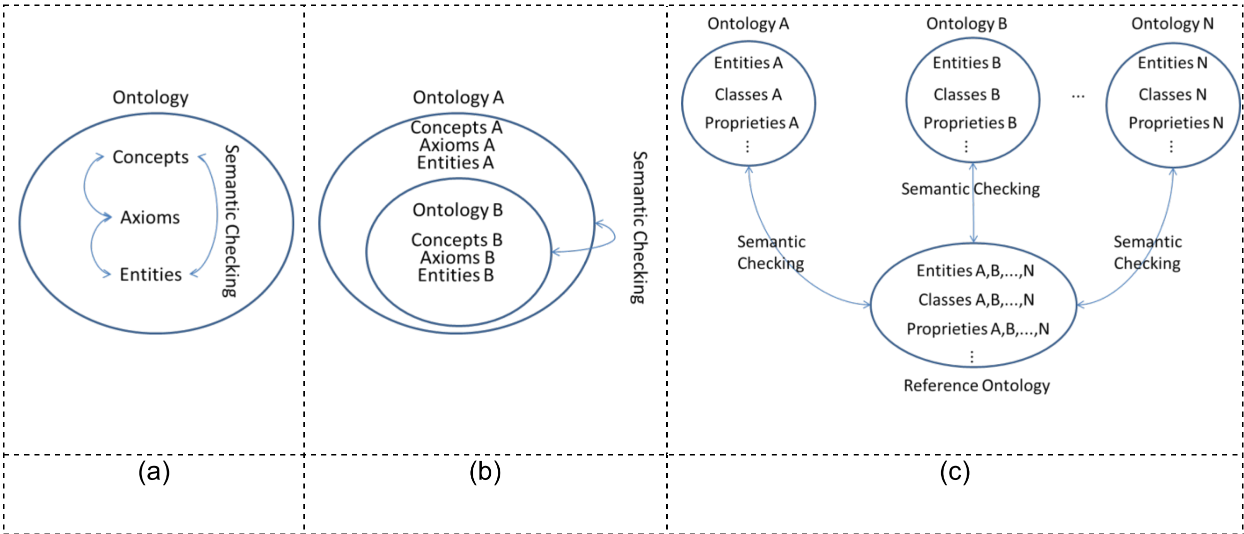


Figure 8.2 – Ontologies Semantic Checking: (a) Single; (b) Composite; (c) Multiple

The proposed framework main purpose is to provide generic guidelines for the se-mantic checking of a KB. The framework is composed of 6 items based on the three types of ontology consistency presented (Table 8.1).

Items 1 and 2 refer to scenarios where only a single ontology is involved. For item 1, a simple reasoning process suffices to verify the structural consistency of the ontology. In addition, item 2, also requires human action. This is because the user needs to create instances of the concepts to test if after running the reasoner, such concepts are well positioned in the ontology.

	Single Ontology	Composite Ontologies	Multiple Ontologies
Structural	1.Automatic reasoning	3.Automatic reasoning; Automatic synchronization	5.Ad hoc synchronization; Automatic reasoning
Conceptual	2.Human action plus automatic reasoning	4.Human action plus automatic reasoning; Automatic synchronization	6.Human action plus automatic reasoning; Ad hoc synchronization

Table 8.1 – Semantic Checking Framework

Items 3 and 4 of the framework denote cases where the KB aggregates various ontologies. Since composite ontologies are composed of two or more ontologies merged together, if a structural change occurs in one of the ontologies, then this change needs to be reflected in all the elements. Thus, item 3 in addition to an automatic reasoning process, an automatic synchronization mechanism is also required for such structures validations. On the other hand, item 4 additionally requires human interaction to the automatic reasoning and synchronization processes. Here, the user also needs to create instances with the same objective mentioned for the item 2. Moreover in this case, the concepts need to be well represented in the merged ontology to avoid repetitions in both ontologies. That is why the synchronization and reasoning are both required.

Finally, items 5 and 6 of the framework are applicable in scenarios where multiple but separate ontologies or a mix of KB types (e.g. ontology and Wiki) are involved. In item 5, besides having an automatic reasoning process, an ad hoc synchronization process is also required, to align the knowledge represented in the various components. This means that any changes that occur in a certain element of the system must also be reflected in all the other components. Since these types of systems can be very complex, knowing the synchronization process facilitates the further semantic checking. In entry 6 it is also needed human intervention, by the same reasons as in the other conceptual checking items. Moreover, an example of this semantic checking kind was used on the MENTOR demonstration (sub-chapter 4.5.3).

To accomplish the communication checking between ontologies it is also needed to know its particular synchronisation process to then execute modifications in one side that could be reflected in the other side. In relation to this, author presented in sub-chapter 6.3.2, MIRAI that is a framework to monitor all the mappings that exist among the several models (ontologies) used by business partners in the same collaborative network. Its implementation prototype through the use of a MAS, is able to control the changes, warning and proposing new mappings, thus preventing inconsistencies (interoperability problems) that could cause a destabilization of the network harmony. Thus, it can be used for this kind of semantic checking.

8.1.2 **Business Validation**

Nowadays, enterprises have began to realize that it is important to stop "playing" by themselves and begin to cooperate or interact with others to be able to answer to customer needs. Collaboration also helps small businesses to join efforts to produce products that can compete with products from large companies [339].

The collaboration models shall focus on semantically enabled peers to facilitate e-collaboration by building added value and mutual trust in diverse European business, cultural and language environments. It means defining business processes and rules for companies and users to participate in a Single European Electronic Market (SEEM). The SEEM intendeds to be an e-Business space where companies of any kind, size and geographic location can work and collaborate without relevant constraints, and supporting the establishment of value-creation relationships while ensuring a high level of security, trust and confidence together with a highly scalable and flexible infrastructure [340].

However, these collaborations arise with problems that are particularly related with the lack of interoperable systems and software applications that manage and contribute to progress in business with each other [1][2]. To succeed in such complex environments, companies need to share their skills and information with each member of the network and this information must be fully understood by all the stakeholders. The problems that these companies face are related with semantic interoperability for efficient enterprise collaborations. Such problem areas (Collaboration Oriented Knowledge and Semantic Interoperability) were identified by the roadmap defined by the community of European reference for EI as one of its major research objectives in the coming years. As a consequence there is a major focus of the current research in supporting the organization of semantics for the administration and exchange of knowledge and information in modern SMEs [16].

The semantic interoperability problems mainly arise because of the different interpretations of the meaning of concepts. Therefore, as proposed in chapter 4, the collaborative construction of a common KB of mutual agreement plays a key role in establishing a common understanding of the semantics of information in communication between companies. Thus, to establish business partnerships it is relevant that companies adapt their KBs, or build a reference so that the harmonization represents the shared knowledge.

However, it is reasonable to say that interoperability between the parties in the same field is still a challenge in progress due to the different systems, and semantic models implemented [341] and especially **because of the expenses in managing inter-enterprise knowledge**.

Business Collaboration and Semantic Interoperability

Collaborative working focuses on the goal to remove the barriers that globalized industries and supply chains face, and to significantly improve communication, knowledge sharing, collaboration and reactivity, to link creative capacities and to reduce costs [342]. Enterprise collaboration is then the "partnering of activities, knowledge and assets by multiple stakeholders in a dynamic environment with the objective of gaining business advantage" [342]. The success of the collaboration models will be measured in terms of turn-around times for locating resources, reduced transactions costs, and increased profit due to costs reductions and new business opportunities as a consequence of operating in larger markets as well as enhanced quality of businesses and services [342].

In addition, the capacity of handling knowledge for business collaboration support and for semantic adaptation of information systems plays an important role in today's collaborative electronic market.

To be able to negotiate between two organizations it is necessary that their representative ontologies have at least parts of them that coincide. Experience shows that differences in the way of seeing a negotiation area undertake the exchange of information between the entities surrounding. This requires that organizations that intend to cooperate, agree on a common ontology or understanding.

Thus the author proposes a methodology to analyse the economic viability of companies in the needed effort to the conception of a common KB in their operational domain area, in order to stimulate interoperability in the business cooperation. These studies endorses the vision that argues that in the future, the capability to adapt the semantics of business information systems will boost the enterprise collaboration providing them new business opportunities.

Economic Viability Methodology

The semantic interoperability resolution is economic analysed as an investment project that an enterprise has to follow. Author created a specific economic viability analyses methodology (Figure 8.3). The proposed methodology was defined based on preparation phases of investment projects defined in [343]. It is divided into seven distinct phases in support of a Semantic Interoperability Resolution Projects (SIRP). It starts from a project concept, through the maturity of the problem, identification, maturation and solution. Some reformulation may occur in the decision or "ex-post" evaluation phases with the aim of improving the proposed objectives of the project.

The first phase is related to the study of variants, which implies a pre-feasibility study of technical variables, as for instance: data exchange standards; type of information systems to be used by companies; time spent with the exchange of information, etc. These studies include technical, organizational, economic and financial variants, taking into account the availability of effective inputs, and conditioning factors resulting in qualitative and quantitative requirements for the resolution of the problems.

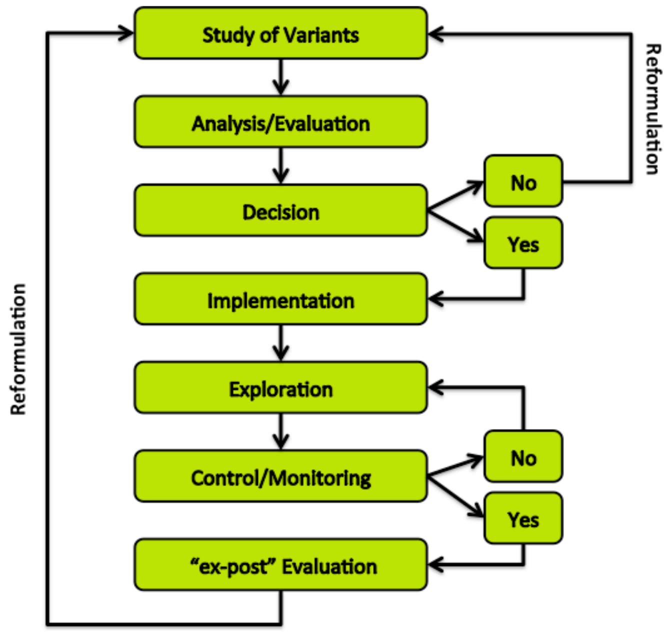


Figure 8.3 – Methodology for Economic Viability of SIRP

The second stage performs analysis and evaluation of the proposed SIRP in the following perspectives: 1) Financial/Business perspective that focuses on funds invested; 2) Economic perspective that includes priorities of the economy; 3) Social perspective that seeks to articulate the analysis of "Welfare" which could define the social groups that benefit most from this SIRP.

These perspectives together with the qualitative and quantitative requirements identified in the previous phase are used for assessment criteria definition. These criteria are then used to assess business evaluation methods to facilitate its choice in the next phase.

The third phase is when it has to be decided if the SIRP will be implemented or not. To support that decision it uses the business evaluation method better scored in the previous phase. Thus, through this method application it proposes a decision, which could be justified by the method itself. However, the nature of the decision is conditioned by the behaviour and operation of the business group that is involved in the process.

If the decision is not favourable, there are two options, the SIRP is abandoned or reformulated by performing a new study, returning in this way to the first phase.

If the decision is favourable, it proceeds to the next phase, the phase of implementation, which includes interventions in various fields, such as:

- Negotiations between customers and suppliers;
- Detailed planning of investments;
- Launch and implementation of an effective management system (the part that involves human resources);
- Implementation of training programs and learning of recruited staff;
- Development of sectorial studies;
- Coordination of projects and installation of equipment;
- Execution of the project;
- Creation of structures for inspection and supervision.

Having success in the implementation phase, it should proceed to the next stages, the phase of exploration followed by the control and monitoring phase. These two phases (exploration and control/monitoring) are inseparable, because since early on, it is necessary to monitor the SIRP in a constant way. Thus, it is necessary to perform periodic analysis/studies, whose results obtained are used to formulate correction schemes to be introduced in the future management actions.

If during the control/monitoring phase it comes to the conclusion that the SIRP can`t be improved, it means that the SIRP reached an optimal point where it is not needed to proceed to improvement actions. In other words, the project stays in the “exploration” phase. However, if the answer is positive, the project can be improved. It makes sense to do an "ex post" evaluation, which allows the realization of adjustments, correcting possible mistakes that occurred in the past, to then reformulate the study of the variants in a new SIRP cycle.

Demonstration Scenario in Aerospace Industry

In today’s context of strong competitiveness, European aircraft, engine and equipment manufacturers are facing greater challenges than ever before. The market demands that more complex products be developed with shorter lead times and more cost effectiveness while using evolving business models aggregating multiple organisations [344].

The collaboration between companies is a key technical challenge in the aerospace product integration process, because usually, different companies use different ICT systems and data stemming from one firm could be used only within that entity and its system.

In the presented aerospace demonstration scenario of the integrator company offices, as show in Figure 8.4 are distributed in France, Germany, UK and Spain. Each office is responsible for some parts in the design and production phases. Virtual 3D models are created for each part. The Digital Mock Ups (DMU) is used to integrate the various model parts to a complete aircraft. Problems normally occur during the integration process [342]. That’s why SMEs in the aeronautics sector face an increasing pressure from their customers to implement ICT applications that support the integration of inter-firm processes [342].

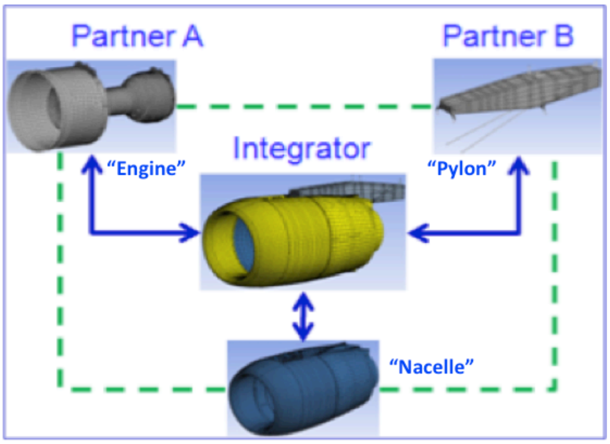


Figure 8.4 – Powerplant Product Integration [345]

In this specific aeronautic case, the principal company awards to several different companies, several distinct products with the ultimate goal of obtaining a unique product, which can be assembled like a puzzle, as illustrated in Figure 8.4. Though, it is necessary to promote the collaboration between companies represented in Figure 8.4. To exchange the necessary information between these companies, both need to know all the specifications and system models about the "pylon", "nacelle" and the engine, so the companies can individually simulate the results, before assembly of the various products [345].

However, there are some constraints, because each company uses a heterogeneous Information Technology (IT) system and do not want to share the proprietary data. To solve this problem, it is necessary to motivate the sharing of some information between these partners, at least to facilitate the integration of the various components in a way that could be understandable by their heterogeneous tools, establishing collaboration and interoperability at several levels (semantic, technical, etc.). One possible solution, as already proposed before, is to use MENTOR to build a reference ontology to be their common KB. Although, it is not easy to push enterprises to follow a methodology as MENTOR without proving to them specific business profits of such involvement. Thus, author proposes the described economic viability methodology to help motivate such enterprises to be involved in these processes as in other similar cases for semantic interoperability problems resolution. In the following it is described this aeronautic case study according to each of the methodology phases proposed.

Phase 1 - Study of Variants

The "Study of Variants" describes the referred scenario mainly the related scenario from a technical point of view. Partner A (engine supplier) and Partner B (pylon supplier) use different CAD systems and information model systems. In addition they also use different nomenclatures and meanings for the same objects. Figure 8.5 presents the points of the interface on the engine and pylon integration. These points represent the minimal information that Partner A and B need to share in order to both simulate their own products and for the integrator partner to do the same.

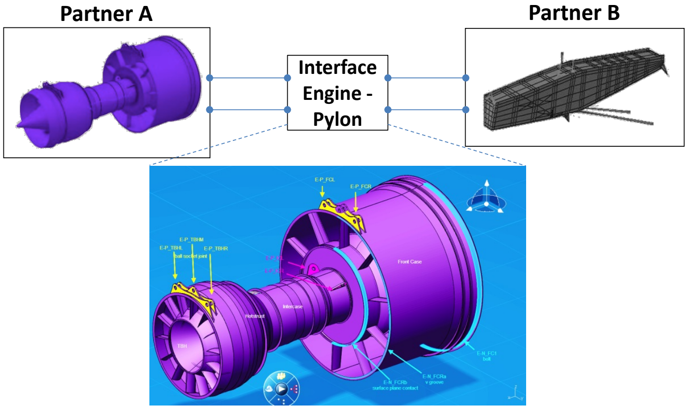


Figure 8.5 – Ports and related publications between engine and pylon [345]

In this study, software engineers identified what is needed to support the definition and specification of sub-systems and components' interactions and related interfaces: as the one presented in Figure 8.5 between engine and pylon. With such work it should be specified the functional, structural, behavioural and modelling parameters of the sub-systems interactions to integrate or reference consistently designed information from the product definition with the associated simulation data. It ensures a complete traceability of the design/simulation information chain [345].

In addition to these specifications, definition is needed to solve the semantic interoperability between the involved partners. To solve this issue MENTOR is a possible solution or process to follow. Thus, the process for building a reference ontology has to be economically validated due to the effort that it requires even for a small set of concepts as shown in Figure 8.5. The main idea identified for the SIRP is to use the business object model as a standard pivot model to exchange engineering data objects supported by MENTOR for its semantic alignment establishment and complemented with a model transformation solution.

Consequently, the communication between partners will be facilitated due to a communication hub that will have an ontology reference with specific semantic mappings for a semantic alignment establishment between them (Figure 8.6), and through a specific model transformation between partners (enterprises) and the communication hub. However this demonstration will be focused on the MENTOR related work.

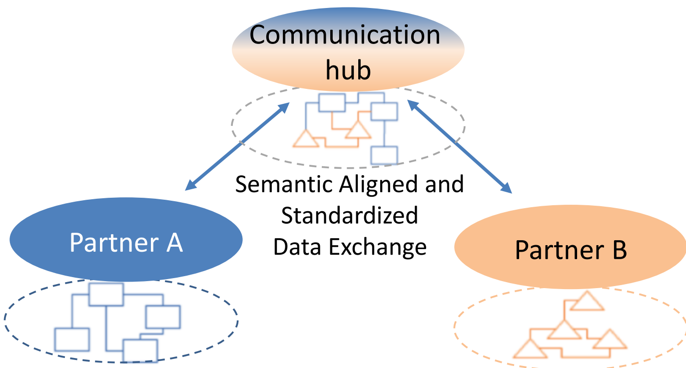


Figure 8.6 – Communication between partners

Phase 2 – Analysis/Evaluation

In the second phase of the economic viability methodology it uses defined assessment criteria to evaluate business evaluation methods. From the technical facts presented before, author defined four criteria. However, as Nekvasil M. stated in [346] an evaluation through criteria are far from forming an exact metric but they can generalize clues for choosing a particular assessment method for a given investment project.

Predictive abilities – This criterion should be based on the desired result and its accuracy and the possibility of collecting data in order to prevent the occurrence of errors, trying to minimize costs if these occur. The company will always choose a method with highest predictive abilities but while this is quite general it will do so under conditions given by the other overall restrictions and costliness of the method [346].

Avoiding Costs – are related to the ex-ante activities stakeholders undertake to prevent or minimize the impact of technical interoperability problems before they occur [347].

Mitigation costs – result from ex-post activities that respond to interoperability problems. These costs are intended to correct problems that have occurred.

Delay costs – arise from interoperability problems that delay the completion of a project or the length of time a facility is not in normal operation [347].

There are many business evaluation methods to assess the investments and each one should be used depending on the different conditions and circumstances of the project and its stakeholders. Since there are no available quantitative values to take in consideration for this SIRP analysis the appropriated economic evaluation methods to be used are mainly the qualitative ones. Thus, the author has chosen the following four assessment methods: Payback, Activity Based Costing, Six Sigma and IT Portfolio Management (ITPM).

Payback – is a method for evaluating the risk, being the most attractive projects, those which allow a recovery of invested capital in less time. This is an important complement of a feasibility study or an evaluation of an investment. Payback gives us the number of time periods required to recover the initial investment. The plain payback time of an investment is the simplest and least adequate but on the other hand mostly uses economic criteria. The biggest gap in this assessment method is that it ignores all effect of the investments after the payback. The limitations of this method (inability to regard changing cash flow over years or the ignorance of the time factor) can of course be recompensed by more sophisticated elaborate approaches of payback time calculation. The core of this method is the time estimation of the monetary return of investment and as such it cannot be used as the main criteria for assessing an IT project. However it can serve well as secondary supporting criteria in the turbulent environment of information technologies; if the estimated time of new release of a product version is shorter than the payback then the investment efficiency is questionable. Thus in the context of IT and especially of a SIRP it makes sense to compare the payback time with the project's operational life time [346].

Activity Based Costing (ABC) – This approach is built upon a costing model that assigns costs to the resource activities to resulting products and services. The emphasis in this approach lies on determining the unit costs of diverse activities by dissolving all the indirect costs into the processes they are tied with [346]. ABC recognizes that the special engineering, special testing, machine setups, and others are activities that cause costs—they cause the company to consume resources. Under ABC, the company will calculate the cost of the resources used in each of these activities. Next, the cost of each of these activities will be assigned only to the products that demanded the activities [348].

Six sigma – is a model of quality management, measures the amount of defects in a given process, it is also possible in a systematic manner how these can be eliminated, approaching the "zero-defects". It is therefore a method that focuses on the prevention of defects by using statistical tools and is applicable to any process within the organization [349].

The basis of Six Sigma calculation is formed by the properties of the normal distribution curve (or the bell curve). In this method we only need to know two variables, the mean and the standard deviation (sigma) to get a whole lot of information about it. The properties of the bell curve are as such that we can know the probability of getting each value within the curve (probability distribution function) and we can also know the probability of getting a value less than or equal to a number of interest within the curve (cumulative distribution function).

For a process to be at Six Sigma level, it needs to have more or less 6 standard deviations within the specification limits in the short term and more or less 4.5 standard deviations within the limits in the long term [350].

IT Portfolio Management – as described by Lih-Bin Oh et al. in [351], “is the combination of tools and methods used to measure, control, and increase the return on both individual IT investments and on an aggregate enterprise level in a desirable manner that meets the organization’s business objectives without exceeding available resources or violating other constraints”. Examples of ITPM would be planned initiatives, projects, and on-going IT services (such as application support). The promise of ITPM is the quantification of previously informal IT efforts, enabling measurement and objective evaluation of investment scenarios. ITPM is a management method for administering a large portfolio of IT projects, applications, processes, services and resources [351].

After having identified the criteria and the methods it is possible to do an analyses to facilitate the choice of the most appropriated evaluation method to this SIRP. Each business evaluation method receives a score (Low, Mid and High), accordingly to its applicability to each of the defined criteria (Table 8.2).

The Payback method received a low score to all criteria because the PRI cannot be used neither for financial analyses of investment nor for comparison of completely different projects and it calculation is not very demanding [346].

	Business Evaluation Method			
Assessment Criterion	Payback	Activity Based Costing	Six Sigma	IT Portfolio Management
Predictive abilities	Low	High	Low	High
Avoiding costs	Low	High	Mid	Mid
Mitigation costs	Low	Mid	Mid	Mid
Delay costs	Low	Low	Mid	Mid

Table 8.2 – Assessment of Business Evaluation Methods

The ABC method involves variables of the diverse activities in the company and doing an approach that lies on determining the unit costs and dissolving all the indirect costs into the processes they are tied with, which is why the score at the predictive abilities and avoiding costs are high. As this method neglects other impacts of investments like in sales, it is an insufficient method to assess the IT investments. That's the reason for a mid score in mitigation costs and low score in delay costs.

Six Sigma model doesn't measure all the variables in a process. That's the motive behind the low score in the predicative abilities. The author ascribes a mid score to the other three criteria because this method denies IT investments projects that have a significant impact on current processes structure and favours projects that support rigidity and are inconsistent with innovation. This is a method for process management monitoring and evaluation that aims to minimize process exceptions and failures.

In the ITPM the author attributes a high score in the predicative abilities because the analysis of this method focuses in different areas in the company that requires some predictability such as: organizational, technical, financial and economic. ITPM method administers a large portfolio of IT projects, applications, processes, services and resources. And a mid score to the other criteria (avoiding, mitigation and delay costs), because the main objective is the mapping of these expenditures, which IT budgets do not usually cover in sufficient detail and therefore do not allow their effective monitoring. It is not sufficient to monitor only financial indicators and is necessary to take into account non-financial aspects.

Phase 3 - Decision

After having the assessment table fulfilled, it is time for the decision stage, where it is defined if there is or is not an implementation viability of the project. Looking at Table 1 we concluded that the most interesting business evaluation method to use in the economic viability analysis for this SIRP is the ITPM method. However, in practice doing an evaluation about a specific project, it is necessary to have high information in all its area's (organizational, technical, economic, and financial) of a company or project. This method may be used with another's quantitative methods with the objective of complementing the study of a SIRP and reinforce the conclusions achieved.

In this paper scenario a set of latent variables or requirements were identified that are related to characteristics of the companies' collaboration establishment as they exchange the necessary information and build a common (reference) KB. Thus, a possible measurement of the ITPM application could result from a qualitative analyses of the SIRP identified variables in relation to the enterprises evaluation areas characteristics, which in these cases are represented by the criteria (predicative abilities; avoiding costs; mitigation costs; delay costs). To have a good result on this business evaluation, a quantification method able to measure the usefulness of participating in a project as the one related to this scenario, is needed to accomplish the ITPM. The purpose of the process is to prove that by implementing the referred solution, the involved enterprise(s) will benefit from it in a direct relation to the mentioned criteria.

Phase 4 - Implementation

This project's goal is to build a common KB of the involved enterprises business to prevent errors in the exchange of semantic data, which are related to predictive abilities criterion. This will minimize avoiding costs since in semantics if you have already a common KB, it will facilitate future semantic alignments. The presented solution intends to reduce mitigation costs, by solving current interoperability problems related to companies' information sharing to facilitate their products simulation. The process for building a common KB intends to follow the MENTOR methodology, which could be a slow process. However, it doesn't require the accomplishment of the process at once. Then enterprises can repeat the MENTOR as often as they need, avoiding in that way delay costs. In conclusion, the appliance of the ITPM considering the presented evaluation criteria can result in a positive decision. Due to its qualitative measurement nature it doesn't result in an exact value for decision. Nevertheless, this SIRP has been implemented in the CRESCENDO project.

Phase 5 - Exploration

This SIRP was not explored yet, because it was only recently implemented. Consequently, it is expected that such project implementation effort will result in the use (Exploration phase of the methodology) by the enterprises involved, which are represented by the partners of this scenario.

Phases 6 and 7 – Control/Monitoring and "ex-post" Evaluation

The phase Control or Monitoring has the objective of performing periodic analysis to the exploring state of the SIRP. Author identified two possible mechanisms to accomplish that: the Business Model Measurement Interoperability Quotient (BIQMM) [352], and maturity models, as for instance MMEI (Maturity Model of Enterprise Interoperability) [353].

The BIQMM was designed for quantitative and qualitative assessment of the interoperability between companies. This analysis is made with the intention of not only assigning scores to how interoperable two companies are but also highlighting the key areas where interoperability needs to be enhanced in the short and medium term, to avoid key bottlenecks to achieve the overall strategic aims of the collaboration [352]. If BIQMM was applied to this scenario, one possible question for assessing interoperability with respect to business semantics would be: Do you and your collaborator use (any new) different terminologies with regards to the business area that you both operate on? If the answer was yes SIRP would proceed to the “ex-post” evaluation, where the problem would be analysed to reformulate the project solution. Then the methodology would go for a new entire cycle. If the answer was no, everything was ok and no improvements would be necessary.

MMEI is intended to be used for the assessment of EI and by the detection of which might need to be improved to meet the needs and ambitions of the enterprise [353]. Thus, it could be applied for the detection of interoperability problems in the solution developed for the maturity evaluation in which the project is. If the project is not at an optimal stage or if it found a specific problem, the methodology proceeds to the next phase, the "ex post" evaluation phase, reshaping the study of the variants, with the aim of improving the previous found problem.

Conclusions

It has been difficult to explain enterprises that share some of this information to build a common harmonized KB could push them to new and dynamic collaborations and consequently new business opportunities. Thus, author intended to provide them, this methodology, to support them to decide, to promote their participation in such SIRP, taking in consideration the natural evolution of the projects. This methodology was demonstrated using an aerospace scenario, which shares the main characteristics of a SIRP. However it presents an absence of quantitative values. That results in a lack of facts or evidence to be used by deciders. However, the proposed methodology is prepared to be adapted accordingly to the economic values or technical requirements available.

For future studies, it is necessary to test the economic viability methodology in others scenarios. The idea is to find other real scenarios that its stakeholders have or could share some quantitative values in order to test the methodology created, using quantitative business evaluation methods as: Return On Investment (ROI); Internal Rate of Return (IRR); Net Present Value (NPV); Payback; etc. The quantitative values referred above are fundamental to obtain results expressed by official current units (money) able to measure costs and consequently to have a more realistic evaluation of a SIRP.

8.2 Research Positioning Assessment

The research conducted for this dissertation is related, as mentioned at the beginning of the document, to EI. However, it aggregates various areas, what conducts to the need to sharply identify its specific area inside EI discipline. To make a correct address of it, it was used the standard that defines such concept through the establishment of a Framework for Enterprise Interoperability (FEI). The purpose of ISO 11354-1:2011 is to specify a FEI that establishes dimensions and viewpoints to address interoperability barriers, their potential solutions, and the relationships between them [354]. Such standard work defined three dimensions of the FEI as illustrated in Figure 8.7. They represent the interoperability barriers areas, the approaches that represent possible solutions to solve the interoperability barriers, and to what concern them (i.e. businesses; processes; services or data).

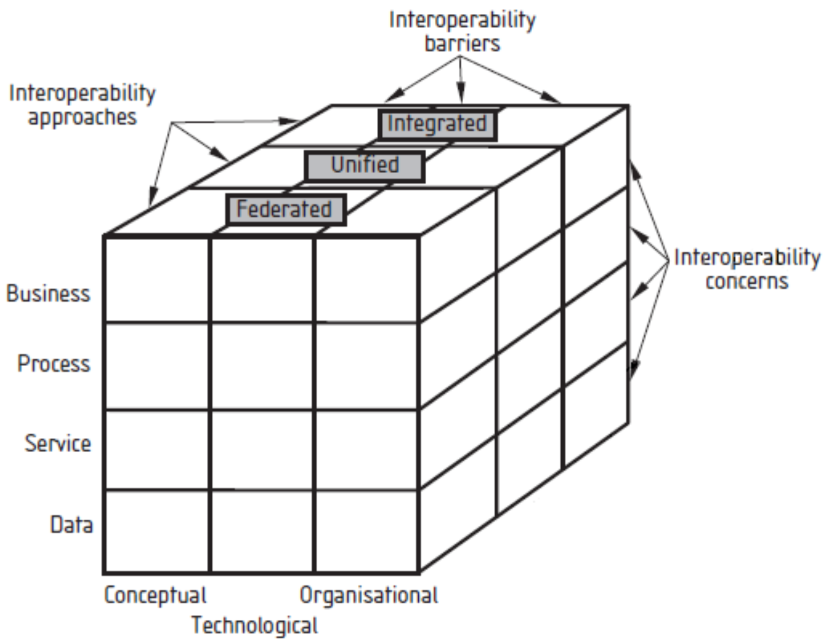


Figure 8.7 – Framework for Enterprise Interoperability Dimensions [354]

In the proposal for the maturity model MMEI, Guédria et al. in [353], presented a diagram aligned with FEI standard to illustrate MMEI coverage, which also introduced other maturity models classified accordingly to this standard dimensions. Since, a maturity model is a framework that describes, for a specific area of interest, a number of levels of sophistication at which activities in such area can be carried out [355][353]. Then, check a maturity model (if existent) that aggregates or coincides with this dissertation should be the correct track to follow. In relation to the FEI presented, this dissertation is mainly focused to the conceptual barrier aspects and concerns to data aggregating federated, unified and integrated approaches. Consequently, due to its relation to these topics (as shown in Figure 8.8) a further analysis of LCIM (Levels of Conceptual Interoperability Model) is presented.

8.2.1 **Levels of Conceptual Interoperability Model**

Turnitsa and Tolk defined in [356], the Knowledge Representation Requirements Model (KRRM), which is a model showing the increasing needs for greater ability to represent knowledge, and also gives the levels of conceptual interoperability that may be reached if the requirements are met. Such conceptual interoperability levels are defined by the LCIM model, which shows the different levels of interoperability that may exist between systems, from technical interoperability through conceptual interoperability [356][357].

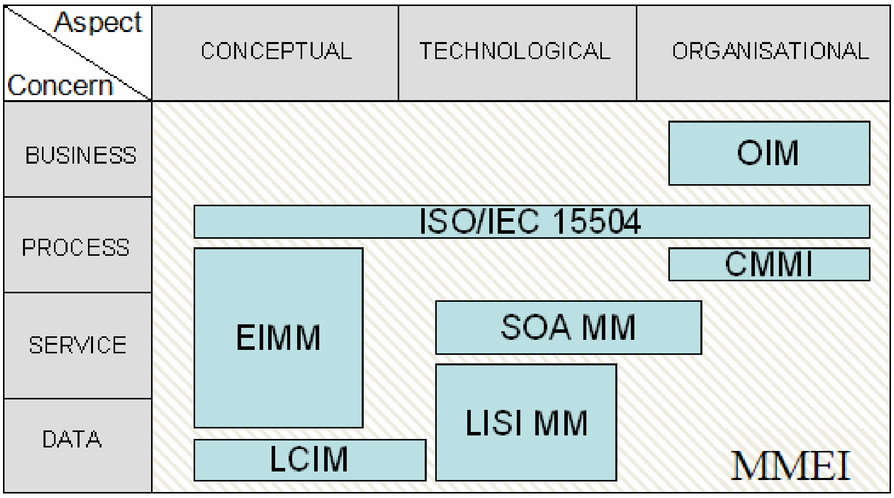


Figure 8.8 – Coverage of maturity models according to EI dimensions [353]

Other authors as Bellinger et al. [48] and Ackoff [253] defined a kind of KR model through knowledge related concepts, establishing a hierarchy between them. As mentioned before, it is composed by DIKW, with “understanding” achieving the transition through the categories. Author proposed the use of the “understanding” as a new level to the mentioned DIKW hierarchy. These hierarchies study the knowledge concept, providing a foundation for exploring the dimensions of knowledge creation and representation. These hierarchies are in an almost accordance to the Turnitsa and Tolk’s KRRM original version. The difference is that KRRM has one more level related to “awareness” before the “understanding” level and do not have the “wisdom” level. Thus, to harmonize the various contributions, author decided to introduce the “wisdom” level on the top of the KRRM original version. Since the idea is to have a LCIM level for each KRRM level, author also introduced the “Philosophical” level to the LCIM model, to be aligned with the KRRM “wisdom” level. This new level is the most intelligent, since it represents the ability of the systems with the ability to put in philosophical probe the previous conceptual interoperability establishments.

This approach resulted in the clarification of the Intelligence Levels in KR requirements and interoperability models. Thus, it is proposed a KRRM version with six levels of requirement, describing what is needed for a system to either represent or accommodate knowledge [356]. The Figure 8.9 shows the KRRM levels on the left, along with the LCIM levels on the right that the KRRM levels correspond to.

As described in chapter 3, research community is battling to reach the “understanding” level of the KR in information systems. Author contributes on this dissertation with a solution to step in forward from the “knowledge” degree to the “understanding” degree by enabling dynamicity to the semantic adaptability of systems. Thus, the work is now in the “awareness” degree, due to its capacity to handle knowledge in a time-sensitive context. This could be related to the dynamic KMa that oLEARCH is able to conduct. It can learn from the users before system administrators do, in a “on the fly” kind process.

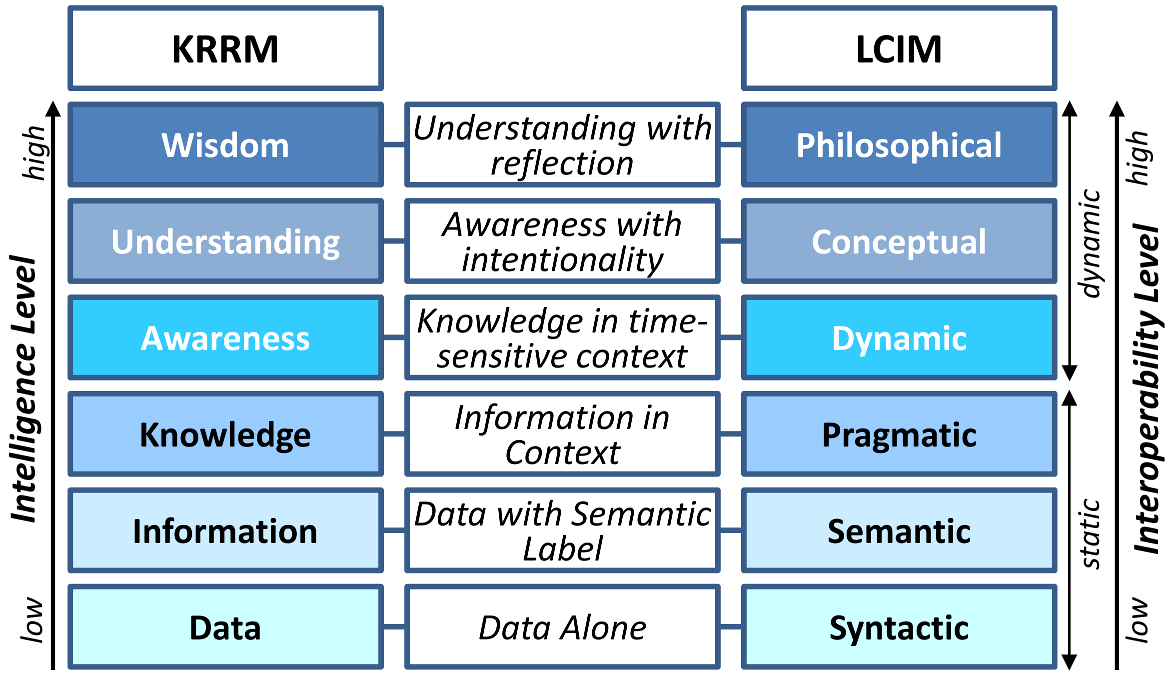


Figure 8.9 – Intelligence Levels in KR Requirements and Interoperability models (based on [356])

8.3 Research Method Assessment

This sub-chapter intends to assess the research method followed by validating the research questions and the hypothesis defined at the beginning of the PhD research work. In additional a description about the integration of the research conducted with other activities is presented.

8.3.1 **Research Questions and Hypothesis Validation**

As stated in the RQ1, the semantic interoperability in information systems can be improved by the establishing of a common ontology able to represent the participants in a specific system. A methodology as the proposed MENTOR, guides a collaborative reference ontology building, defining at the same time, semantic mappings from the participants' proprietary KB to the new ontology and vice-versa. This allows enterprises to keep their internal nomenclatures avoiding change management in their processes. MENTOR has the lexicon settlement and the reference ontology building phases. The first helps on the agreement of the concepts used in the domain and the second to the building of a new structure able to represent this common KR of an enterprise group. The process used to build this reference ontology uses harmonisation processes that are related to the taxonomy and to the properties and rules definition of an ontology, keeping at the same time old ontologies and people views working in an harmony to this new ontology. All of this effort will enable efficient communications between the various participants through transformations, as the ones presented at the MENTOR demonstration (sub-chapter 4.5.3).

This approach is able to improve the semantic interoperability of information systems, thus, it answers to the RQ1 following the H2. The semantic mappings mentioned have the role to allow the communication and to establish the understanding between the various concepts owners' views. In a similarity to the synapses do in the brain, semantic mappings are used to verify the importance (weight) of a specific concept connection, enabling its related ontologies to learn. Consequently they enable its system to increase its intelligence level. These statements answer RQ2. The harmonisation of ontologies enabled the establishment of these mappings, then it could be concluded that by inheritance harmonisation of KR also contributes to the advance of the integrated information system's intelligence. This answers RQ3.

The ontology learning approach proposed (oLEARCH) uses specific statistics and a machine learning approach associated to the mappings existent between ontologies. oLEARCH learns from users feedbacks. The analysis of users' usage patterns helps on the proposition of new concepts to the information system, in a kind of KMa approach. Such process is aligned with H1. Moreover, the fact of having a system able to dynamically react to new introduced concepts and learning from them, demonstrates the appearance of a new generation of intelligent systems, reaching the "awareness", as defined in Figure 8.9, and this is aligned with the H3.

8.3.2 **Integration with other Research Activities**

The integration of a PhD research with other research activities conducts to ideas discussions, share of experiences and creation of collaborative networks with peers.

As stated previously, the author participated on several research projects before to his involvement in this PhD work. It helped him to create some critics mass and to create some social networks with other researchers on the area. Thus, in the following it is presented a list of research projects that the author participated. These projects are presented with a short description of the added value that such participation resulted, directly related to this PhD research subject:

- IDEAS – In this project author contributed to the building of a glossary about interoperability. To such objective author built as well a prototype to support such building process based on protégé and in web services. A paper related to this, with the title "Ontology-based Framework for Enhanced Interoperability in Networked Industrial Environments" was published in the conference INCOM 2004[39].

- INTEROP – In this project, the author contributed with ideas for the semantic enrichment of models. A paper related to this, with the title "Annotation for Enterprise Information Management Traceability" was published in the conference ASME 2007[40], which conducted to the former view of the mediator KB.

- SMART-fm[41] – In this project, the author participated in the creation of an ontology based on the harmonization of existent ones on the furniture domain. From this work two publications were produced focused on a methodology for building ontologies based on the harmonization approach, to which the author contributed for, even after the project ending: 1) "MENTOR – A Methodology for Enterprise Reference Ontology Development", In the : 2008 4th International

[39] INCOM 2004 – 11th IFAC (International Federation of Automatic Control) Symposium on Information Control Problems in Manufacturing, Salvador, Brazil, 2004
[40] International Design Engineering Technical Conferences & Computers and Information in Engineering Conference (IDETC/CIE 2007), Las Vegas.
[41] IST research project – info available in the web (retrieved at August 2011) - http://www.ist-world.org/ProjectDetails.aspx?ProjectId=85ca05c5b41f48ab869ad8494282eecb

IEEE Conference on Intelligent Systems, Varna, Bulgaria; and 2) "MENTOR: an enabler for interoperable intelligent systems", In the International Journal of General Systems, in 2010. In additional, a prototype related to MENTOR was also developed.

- INNOVAFUN[5] – In this project, the author participated in the definition of a solution to the Semantic Enrichment of Standard-based Electronic Catalogues (AP236[42]) of the furniture domain. A paper related to this, with the title "Semantic Enrichment of Standard-based Electronic Catalogues" was published in the conference IFAC INCOM 2009, and in the Journal of Intelligent Manufacturing a publication was also made with the title " Knowledge framework for intelligent manufacturing systems".

- iSURF[43] – In this project the author participated in the evaluation process of the prototype developed. From this integration work, the author defined a methodology to accomplish evaluations on prototype pilots' execution. Such evaluation methodology after some adaptations was used to evaluate and validate his work on this dissertation.

- CoSpaces - In this project the author participated on the organization and development of the training courses and programmes. Such work produced a knowledge model of the training developed that was used in the prototype also developed by him, which is able to create on the fly adaptable training for trainees intended topics and profiles. Two papers were published related to this work: 1) "A Training Curriculum in Collaboration for Engineering Management" published in the IEMC-Europe 2008[44]; and 2) "The CoSpaces Training System" published in the ICE2008[45] conference.

- CRESCENDO[46] – In this project the author has been participated by contributing with a proof of concept that like this dissertation is focused on the semantic interoperability establishment in the knowledge collaboration paradigm in study on the domain area of the project - behavioural digital aircraft. Such work resulted on a prototype (oLEARCH) and in a published paper: "Knowledge Based Methodology Supporting Interoperability Increase in Manufacture Domain" published at ASME IMECE 2011[47]. In additional on this project his work was also related to the training organization and in a glossary building process, which conducted to the publication of the paper: "An Inter-Organisational approach to Industrial e-Training", published at CKM 2010 – IMETI 2010[48].

[42] Part 236: Application Protocol For Furniture Catalog and Interior Design. The standard was published in 2006 as ISO 10303-236:2006(E)
[43] Research project: An Interoperability Service Utility for Collaborative Supply Chain Planning across Multiple Domains Supported by RFID Devices – website: http://www.srdc.com.tr/isurf/ (accessed at August 2011).
[44] International Engineering Management Conference, 28-30 Jun 2008, Estoril, Portugal.
[45] 14th International Conference on Concurrent Enterprising, 23-25 Jun 2008, Lisbon, Portugal.
[46] http://www.crescendo-fp7.eu/ (accessed at August 2011).
[47] ASME International Mechanical Engineering Congress and Exposition IMECE 2011, USA
[48] Collaborative Knowledge Management - CKM 2010: Inter-Organizational Collaboration, Collaborative Software Development, and Knowledge Management (I) - Invited Session of the 3rd International Multi-Conference on Engineering and Technological Innovation: IMETI 2010 – Orlando, USA.

- InterOP-VLab[49] – The author is a member of the InterOP-VLab committee for the task force 6: Ontology for EI[50]. His main role has been to research ontology based advances/application for EI, in cooperation with peers (also members of the i-VLab).

In additional the author has published other papers and has participated in other activities like:

- Synergies with MSc research works that are related to this PhD topic: 1) "Methodology for Collaborative Enterprise Reference Ontology Building" – presented at FCT/UNL in 2011; 2) "Customisable e-Training Programmes based on Trainees Profiles" – presented at FCT/UNL in 2011; 3) "Ontologies may Learn" – presented at FCT/UNL in 2011; and 4) "Metodologia para análise da viabilidade económica das empresas na adaptação semântica dos sistemas de informação, como estímulo à interoperabilidade na colaboração empresarial" – to be presented at Universidade Lusíada in 2012.

- Conferences organisation (e.g. CE2003; IESA 2007 ICE2008 and DoCEIS 2010).

- Seminars presenter: 1) "Fundamentals on Collaborative Working" – At VRAP 2007 (3rd International Conference on Advanced Research in Virtual and Rapid Prototyping 24-29 September 2007; 2) "Customisable e-Training Programmes" – At Cava 2011 III Congreso Internacional de Ambientes Virtuales de Aprendizaje Adaptativos y Acessibles – 2,3 and 4 November 2011 – Bogotá, Colombia.

8.4 Validation methodology

For the validation of the overall research work conducted it was executed an evaluation process based on a self-assessment questionnaires and in performance indicators. Both of the previous mechanisms are defined following a recognised evaluation framework.

8.4.1 DECIDE Evaluation Framework

Evaluation provides feedback for the evaluator or the company in question, but it also provides important feedback to the participant/team in terms of their progress on a particular task. Evaluation can highlight any strengths and weaknesses of a system (e.g. in terms of usability), and end-user adaptability and expectation can be observed and assessed on predetermined criteria [359].

The evaluation has to be carried out mainly at the end phase of the work. This can allow identify key issues and further user requirements, from which author can work to bridge any identified gaps.

In keeping with Preece et al. in [359] 'DECIDE' framework to guide evaluation, it will be followed the next checklist:

1- **D**etermine the overall goals that the evaluation addresses.

2- **E**xplore the specific questions to be answered.

[49] The Virtual Laboratory for Enterprise Interoperability - http://www.interop-vlab.eu/ (accessed at August 2011)
[50] Task Force 6: Ontology for EI available at - http://interop-vlab.eu/the-scientific-activities/i-vlab-task-forces-descriptions/task-force-6-ontology (accessed at August 2011).

3- **C**hoose the evaluation approach and methods which are appropriate to answer the questions

4- **I**dentify any practical issues which may impact on the evaluation methods used, for example, selecting participants.

5- **D**ecide how to deal with any ethical issues with regards to the evaluation process, for example, privacy and confidentiality issues.

6- **E**valuate the research work, collect and analyse the data and present the finding results in appropriate and clear manner.

Determine the goals

What are the high-level goals of the evaluation? Who wants it and why? What are the user needs? An evaluation could have different goals. An evaluation helps to determine the best metaphor for a conceptual design, or to fine-tune an interface, or to examine how technology/research improvements should be done [359].

The PhD work evaluation follows two kinds of objectives: Research Domain objectives and Regular PhD Overall objectives. Regular PhD Overall objectives' purpose is to evaluate the accuracy of the PhD work performed concerning the overall objectives defined. Research Domain objectives rationale is to warrant that the overall work is compliant with the research community that in this case is related to EI.

Explore the questions

In order to make goals operational, questions that must be answered to satisfy them have to be identified [359].

A set of questions is used to characterize the way the assessment or achievement of a specific goal is going to be performed based on some characterizing model. Questions try to characterize the object of measurement (product, process, resource) with respect to a selected quality issue and to determine its quality from the selected viewpoint [360]. Thus, for each of the presented objectives a set of questions were defined. By answering them, a quantitative answer will be provided to be used to analyse and interpret the results of the evaluation. This process is based on the Goal-Question-Metric (GQM) approach.

GQM is a top-down approach to establish a goal-driven measurement system for software development, in that the team starts with organizational goals, defines measurement goals, poses questions to address the goals, and identifies metrics that provide answers to the questions. GQM defines a measurement model on three levels as illustrated in the figure below [360]:

The insertion of this approach to inside of the DECIDE framework will help to relate three points of its checklist: "Determine the goals"; "Explore the questions"; and "Evaluate, interpret, and present the data"; which are directly connected to goals, questions and metric respectively.

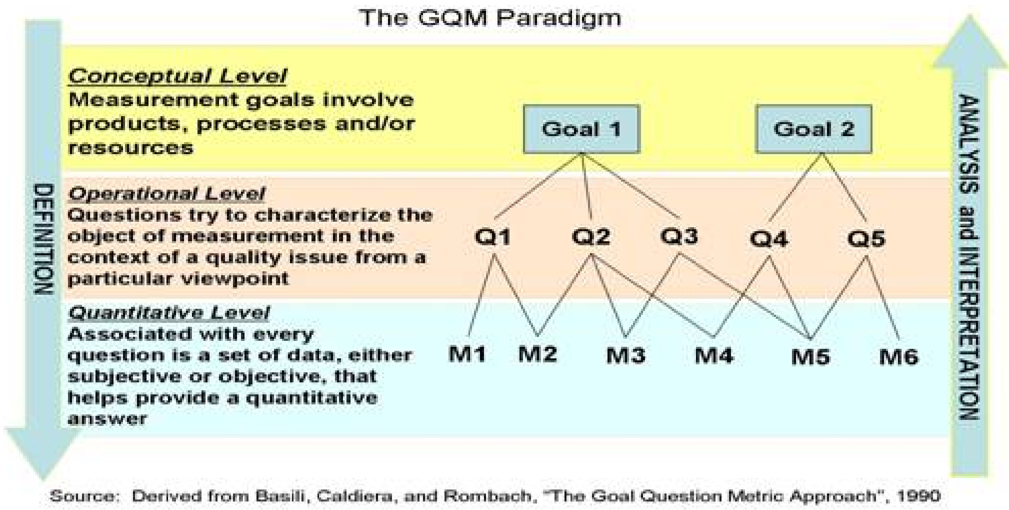


Figure 8.10 - The GQM Approach [360]

Choose the evaluation paradigm and techniques

Having identified the goals and main questions, the next step is to choose the evaluation paradigm and techniques. The evaluation paradigm determines the kinds of techniques that are used. Combinations of techniques can be used to obtain different perspectives. Each type of data tells the story from a different point of view [359].

This evaluation uses three techniques: GQM approach, through a self-assessment qualitative evaluation questionnaire and the Performance Indicators Mechanism.

The GQM is introduced to help on the evaluation measurement definition, specially warranting that all the research objectives were addressed.

Performance Indicators (PIs) or Key Performance Indicators are quantifiable performance measurements used to define success factors and measure progress toward the achievement of specific goals [361]. The PIs mechanism addresses in this case research characteristics as an orientation purpose, helping on guiding the PhD work to reach its own overall research purpose.

PI ID	Performance Indicators (PIs)	Expected Measure
PI1	Number of training sessions delivered about semantic interoperability	2
PI2	Direct contribution to European research projects	2
PI3	Networking establishment / participation in a virtual lab	1
PI4	Publications in Journals	2
PI5	Publications with foreign researchers	3
PI6	Publications with other Portuguese Institutions Researchers	3
PI7	Other Publications (not included on the above)	3
PI8	Collaboration with MSc research works	3

Table 8.3 – Research Performance Indicators

Identify the practical issues

There are many practical issues to consider when doing any kind of evaluation and it is important to identify them before starting. Some issues that should be considered include users/stakeholders, facilities and equipment, schedules and budgets, and evaluators' expertise. Depending on the availability of resources, compromises may involve adapting or substituting techniques [359].

In this case, no practical issues were identified.

Decide how to deal with the ethical issues

Personal records containing details about health, employment, education, financial status, and where participants live should be confidential. Similarly, it should not be possible to identify individuals from comments written in reports [359].

On this specific case no ethical issues were identified.

Evaluate, interpret, and present the data

Choosing the evaluation paradigm and techniques to answer the questions that satisfy the evaluation goal is an important step. Decisions are also needed about what data to collect, how to analyse it, and how to present the findings [359].

In order to evaluate the PhD work it is needed to specify its evaluation criteria to be used. The criteria have to address all the topics identified as objectives, following the GQM approach and in this case, it is attributed an identical importance to all the evaluated topics.

8.4.2 Validation and Evaluation Results Criteria

The validation of the work performed is qualitatively evaluated, in the sense that from the questionnaires results and performance indicators is possible to state if the result it was positive or negative. However the PI will be able to provide a quantitative result. These will represent the percentage of PIs had a success result.

The Performance Indicators Evaluation Results (PIER) are be the number of successful performance Indicators results divided by all performance indicators.

$PIER = PI_{success} / PI_n$

	Poor	Fair	Good	Excellent
PIER	[0 .. 0.25]	]0.25 .. 0.70]	]0.70 .. 0.85]	]0.85 .. 1]

Table 8.4 – PIER Levels

The qualitative evaluation result of each question has four possible qualitative values: Poor; Fair; Good; and Excellent.

8.4.3 Overall Validation Implementation and Evaluation

The validation implementation follows five different categories (goals) that are presented in the following sections.

Research organisation and integration Validation

The research validation addresses the integration of the PhD work with other research activities and relays in the research method confirmation, namely the research questions and hypothesis. Thus the questions are:

1. There are any other relevant research activities conducted during this PhD, integrated with this dissertation work?

As presented in sub-chapter 8.3.2, there are three types of research activities to which author participated on: organisation of conferences; membership of certain research committees, seminars presenter, and participation in international research projects. The answer to this question is yes with a level of **"excellent"** because of the diversity of activities and to the frequency number of executions or contributions accomplished in each type (e.g. participation in organisation of 4 conferences).

2. Are the research questions proposed at the begging of this work (sub-chapter 1.2.1) answered? And such answers were aligned with the hypothesis defined (sub-chapter 1.2.4)?

As presented in sub-chapter 8.3.1 the research questions were answered following the proposed hypothesis. Thus the answer is yes with a level of "**excellent**".

Scientific Position Validation

The scientific position validation checks if the author's research is aligned to the overall research community in the domain. This could be done by two ways: 1) check if the author's contributes to a recognised issue, thus already identified by the community; and 2) check if the author has its own work classified in relation to others work (e.g. standards), having on that way a recognition of his frame (position) in the overall research community of the area. Thus the questions are:

3. Is this dissertation research aligned with the recognised issues of the research community in the EI domain?

Yes. This dissertation research is aligned with the EI roadmap. It defined four GC related to: 1) Interoperability Service Utility; 2) Future Internet and Enterprise Systems; 3) Knowledge-Oriented Collaboration and Semantic Interoperability; and 4) EI Science Base. The work is focused in the third GC contributing directly to its major points that are: semantics organisation, management and knowledge and information exchange within the objective of facilitating enterprise collaboration to exploit new products, solutions or services opportunities. All of the work presented in this dissertation is around these themes. It has the objective to facilitate formation of collaborative consortia to exploit product opportunities and especially to the use of such knowledge to support semantic interoperability of information systems. These points can be easily extracted from the MENTOR, which is a methodology to support collaboration in a common KR development to further exploit to business opportunities like in the sell of products, as presented in the oLEARCH prototype scenario. In additional author also contributed to the ISU GC with the publication [283], which demonstrates the MENTOR application in the support to ISU. Even not contributing directly to the GC2 all the prototype developments were made taking in consideration the ideas stated on these point which calls for solutions should be, e.g., open and adaptive. Finally, some of the research results related to training could be used to support on of the goals of the GC4, which intended to "provide an infrastructure to support the diffusion of ideas, education, research, and training" [16]. Thus the answer is yes with a level of "**excellent**".

4. Was this dissertation research positioned in relation to any recognised research work?

Yes. This dissertation research contribution is mainly concerned to the data issues of conceptual interoperability barrier, as have been defined by the ISO 11354-1:2011 work. More detail is presented in sub-chapter 8.2. In additional, in this research it was used a big list of scientific publications that is presented in the references section. However, for a better analysis they are mostly from five categories or types: 1) 40% from journals and books; 2) 22% from conferences publications; 3) 7% from technical documents (e.g. standards; technical reports; research project deliverables); 4) 26% from web sites; and 5) 5% from the author own publications. Thus the answer is yes with a level of "**excellent**".

5. What's the main difference from your work in relation to others in the same area of research and in what level of excellence do you classify it?

The main difference is that the other existent works in the area are focused in solving mainly atomic specific cases or objectives. One example that demonstrates the applicability of the semantic adaptation for a specific case is on the case of the project COntext INterchange (COIN), which through a mediator it has the objective to create the necessary conversion programs from a small number of conversion rules. Thus, it can automatically generate the over 22,000 conversion programs needed to enable semantic interoperability using only six parametizable conversion rules [358]. This dissertation intends to provide a solution for any domain that requires semantic adaptation proposing the building of a common KB to facilitate mainly SME collaboration in an interoperable way, without needed internal change management, and by the difference, also thinking in the knowledge transfer to their assets enabling evolution of the people but also of the information system. This is the result by following SECI model, which by itself has the foundations to the self-transcending dimension of the knowledge. Thus I classify this work as an **"excellent"**.

Scientific Publications Validation

The scientific publications validation is related to the assessment of the impact that research publications made have. Thus the question is:

6. How can you characterise the impact of your publications?

In total, 31 publications (Annex 10.3), including 3 in relevant scientific journals (with SCI impact factor), have resulted from the research developed. This represents an "h_index" value to 7 and, "g_index" to 9, summing a total of 109 citations. For a PhD student is an excellent value. Thus the answer is "**excellent"**.

Industrial Validation

The industrial validation relates to the coverage of domains that are addressed in this research work, and to the transfer of results to the industry. The questions to answer are:

7. Was various industrial use cases or scenarios used for testing or demonstration in this research work?

It was addressed 4 different case studies from 5 different industry domains. MENTOR was tested using a “bolt” scenario in the mechanical area. The knowledge architecture for multicultural manufacturing systems used a furniture scenario (the “leather couch” case). Then oLEARCH uses both previous scenarios. However in the Customisable e-Training implementation it was addressed training contents from ICT to other three different industries (automotive; aerospace and construction). Finally a business validation addresses an aerospace case. They were used for different components tests or demonstrations. Thus the answer is yes with a level “**excellent”**.

8. Were any of the results of this work transferred to (be used in) industry?

Until the moment to my knowledge none of the results are in a current use in industry. However, CADEF is a tool that uses the semantic enrichment of standard product catalogues presented in this dissertation, and it has been tested in a direct contact with furniture industry, but to my knowledge is not in a regular use yet. The customisable e-Training approach is actually available at Internet to want receive training related to collaborative working. It has particular contents related to the automotive, aerospace and construction industries, but no confirmation that it has been used so far by such industries. Finally, MENTOR even already have been referenced by external researchers in their works and used in knowledge organisation in a few international research projects it was not used yet in an industrial environment. However, author have discussed in his research community how such methodology could be pushed to organise knowledge in enterprises, when due to its required collaborative discussions, it requires a big initial effort in terms of time and consequently of money. Two types of answers have been received: 1) enterprises will not enrol themselves on that except if profits are clearly defined or if they receive sponsorships for that; 2) enterprises will enrol if clearly profits are defined but specially because they do not need to share all the knowledge (neither they want) but only the extreme necessary, thus, time and money efforts decrease a lot. In a pursuit to motivate enterprises to enrol in processes as MENTOR, author defined an economic viability methodology. It analyses the economic viability of semantic adaptation of the information systems, as stimuli to the interoperability resolution and to the business collaboration.

Even not having any of the results being actually in use by industry, especially due to the points presented before, the answer to this point is **“good”**.

Technical Validation

The technical validation is related to the proposed implementations or developments and their relation to the main conceptual contributions or research objectives. The question related to this goal is:

9. Are the developed implementations (e.g. prototypes) able to demonstrate the feasibility of the

DSA framework?

As described in sub-chapter 8.1 the various implementations presented in chapters 4, 5, 6 and 7 intends to demonstrate applicability of the ideas presented in the DSA framework description. Three different elements were developed related to the SIE system and even not tested integrated they establish a set of functionalities understandable as feasible to work together because of the existent of common KRE between them. As an example, reference ontology and thesaurus are built in MENTOR methodology and MO is also used to represent some semantic mismatches found, then oLEARCH uses MO and the reference ontology, and the multicultural architecture proposed uses thesaurus. All of them are able to work sharing such elements, thus, their specific functionalities are possible of being integrated. However EKT system is demonstrated by the customisable e-Training approach. It accomplishes all the stages identified as needed for the EKT system. The missing point is its integration with the SIE system. That is something that is expected to be accomplished in the near future. The first idea is to automatically build training contents related to the semantic adaptations ocurred. With this will be possible to transfer knowledge to people accomplishing an entire cycle at the self-enhancing dimension. However all the methodologies and prototypes defined for the accomplishment of the EKT system stages still be aligned with the rest of the prototypes. Normally e-Training materials use a glossary to establish a common lexicon in the trainees. Glossary is a product from the MENTOR development, thus there is already modules foreseen that will be integrated in the near future.

Due to the lack of an integration test between EKT and SIE systems developed elements, the answer to this point is **“good”**.

Evaluation Results

Concerning the PIs, it was reached a complete success result. Thus it was reached 8 PIs values from a total of 8 (Table 8.5). Thus PIER = 8/ 8 = 1.00 that means an “excellent” result. To confirm this there is in annex 10.3 a list of the author publications.

PI ID	Performance Indicators (PIs)	Expected Measure	Result
PI1	Number of training sessions delivered about semantic interoperability	2	3 (one per year)
PI2	Direct contribution to European research projects	2	3
PI3	Networking establishment / participation in a virtual lab	1	1
PI4	Publications in Journals	2	3
PI5	Publications with foreign researchers	3	7
PI6	Publications with other Portuguese Institutions Researchers	3	7
PI7	Other Publications (not included on the above)	3	13
PI8	Collaboration with MSc research works	3	4

Table 8.5 – Research Performance Indicators Results

Concerning the self-assessment questions it was obtained 7 “excellent” and 2 “good” results, which represents an “excellent” work.

9 Final Considerations and Future Work

The research work conducted proposes a DSA framework that is based on the capability of handing knowledge, following as guideline the knowledge creation model proposed by Nonaka and Takeuchi. Such model presents how knowledge is converted from an individual to an inter-organisational level and vice-versa, in a self-enhancing dimension. In additional, the framework was also co-related to SECI, which establishes the foundation to the knowledge self-transcending dimension. This dimension is related to the capacity of generate new (i.e. tacit not-yet-embodied) knowledge.

DSA framework is composed by two abstract systems: 1) SIE that has the role to accomplish the knowledge conversion from an individual to an inter-organizational level, and 2) EKT that has the role to accomplish the knowledge conversion in the opposite direction.

SIE system is rooted by the knowledge life cycle, which objective is to acquire, model, use and maintain knowledge in a dynamic basis. The knowledge is handled in a kind of a knowledge creation spiral, until entirely represents the participant organizations.

The EKT system is rooted by an ISD approach that has the objective to establish activities, actions, processes and operations to be used as a guide to a strategy and plan definition of the work needed to conceive, develop and implement the training activities required for an effective knowledge transfer. The evolution of the knowledge here is from an inter-organisational level to an individual.

The accomplishment of this process will enable the creation of tacit not-yet-embodied knowledge by an individual, which then, through DSA, it will be integrated in the previous settled knowledge of the system. The vision is that in the future all this cycle could be done entirely by the system, which by that would result in a system able to handle semantic adaptation as humans do. For now, the objective is to try to advance the systems response to the most intelligent level possible. This dissertation concluded that by this framework implementation, systems are able to reach the “awareness” level of intelligence that means that they are able to handle knowledge in a time-sensitive context. Thus, they are able to react dynamically to certain semantic adaptations.

This is the conceptual contribution of this dissertation. It is aligned with the Aaronson’s [362] definition for the conceptual contribution “along the way”, which is an innovative way of modelling, looking at, or manipulating a known object or problem, including establishing a new connection between known objects or problems. Thus, the objective of this dissertation was to produce enough statements, implementations and analysis about the research results that once understood, the conceptual aspects of the work could then be viewed as obvious, which means that they simply have become fully incorporated in the worldview of an expert [362].

SIE system was prototyped by the development of three components. MENTOR is a methodology developed for the collaborative building of an ontology, to represent a common view of a domain, for an enterprises’ organization, requiring a minimum of changes at intra-enterprise level. A multicultural knowledge-based architecture was defined, with the objective to facilitate the complementation of multilinguism issues, in information systems, and by introducing the semantic enrichment capability of

KRE. In additional, oLEARCH is a product searching prototype able to learn from users usage patterns. It establishes the dynamism to the KMa.

EKT system was prototyped by a customisable e-Training service supported by a structured methodology for training courses development. The training courses development methodology has the objective to organise the training implementation to facilitate further training KR. Such knowledge, through the use of ontologies is then used by the customisable e-Training service, to generate courses adapted to a specific purpose or profile.

DSA framework was also accomplished by a semantic checking framework to validate the execution of semantic adaptations, and by an economic viability methodology to facilitate future deployments of these solutions kind in the industry.

DSA definition and its implemented prototypes were developed to answer the research questions proposed in accordance to the hypothesis defined.

Added value contribution of this work

The real added value contribution of this research work is mainly related to the conceptual appliance of the Nonaka`s and Takeuchi´s SOKC and SECI models to the semantic adaptability of systems, which resulted in the creation of the DSA framework. From such framework definition resulted in the generation of other added values: MENTOR, the multicultural knowledge-based architecture and OLEARCH, which by their specificities are able to complement each other in the ability to handle knowledge in an entire life-cycle.

Moreover, the training courses methodology and the customisable e-Training approach which implements the ability to systems make the knowledge transfer to enterprise people, are also other added values from this work. All of these prototypes have been tested in international research projects, independently to the domain resulting in their presentation in scientific publications, which already received some worldwide citations.

Finally, the economic viability and the validation methodologies are both research results that support the semantic interoperability solutions implementation in industry. The validation methodology was already tested in a textile industry with success results.

Future Work

As a future work, author intends to accomplish all the prototypes developed and finalise the SIE and EKT integration, to then better technically analyse an entire DSA framework implementation.

However, the main idea of future work is to continue the research about semantic adaptability of systems, having as pillars the main outputs presented on this dissertation. Such research pillars are: the collaborative building of ontologies; the multicultural issues in the information systems; the ontology learning approach handling uncertainty knowledge; the training implementation approaches especially to its related knowledge handling and transfer; the semantic checking processes; and finally the business validation of semantic adaptation projects.

10 Annexes

10.1 Language Independent Meta-Model (LIMM)

LIMM serves as an abstract interface on top of enterprises' information models. Through its usage, becomes possible to abstract the technology and implementation details associated with the different modelling languages, and thus, enlarge the scope of users involved in a traditional mapping definition activity. Having manager and domains experts involved in this process increases the quality of the mappings that will enable interoperable relationships. In comparison to most modelling languages, it is intended to enable as little loss of expressiveness as possible, but at the same time, be simple and generic to support multiple language mappings.

Also, LIMM resemblances with ISO/IEC (2004) standard are not by fortuity. This abstract interface was based on the standard's foundations and concepts in order to give support to mechanisms for enabling global data interchange, particularly across application areas. A bridge between major LIMM concepts and ISO/IEC 11179 can be made, e.g. the standard's "Entity", "Property" and "Representation" concepts correspond to LIMM's "Entity_Concept", "Property" and "Representation" constructs, respectively. The language independent meta-model proposed is described as an UML class diagram in Figure 4.15. It was designed with an UML tool, since UML class diagrams are a popular form to visualise the meta-model, enabling lossless integration with MOF XMI model which is compatible with most of the model management technologies for defining and executing transformations (e.g. ATL).

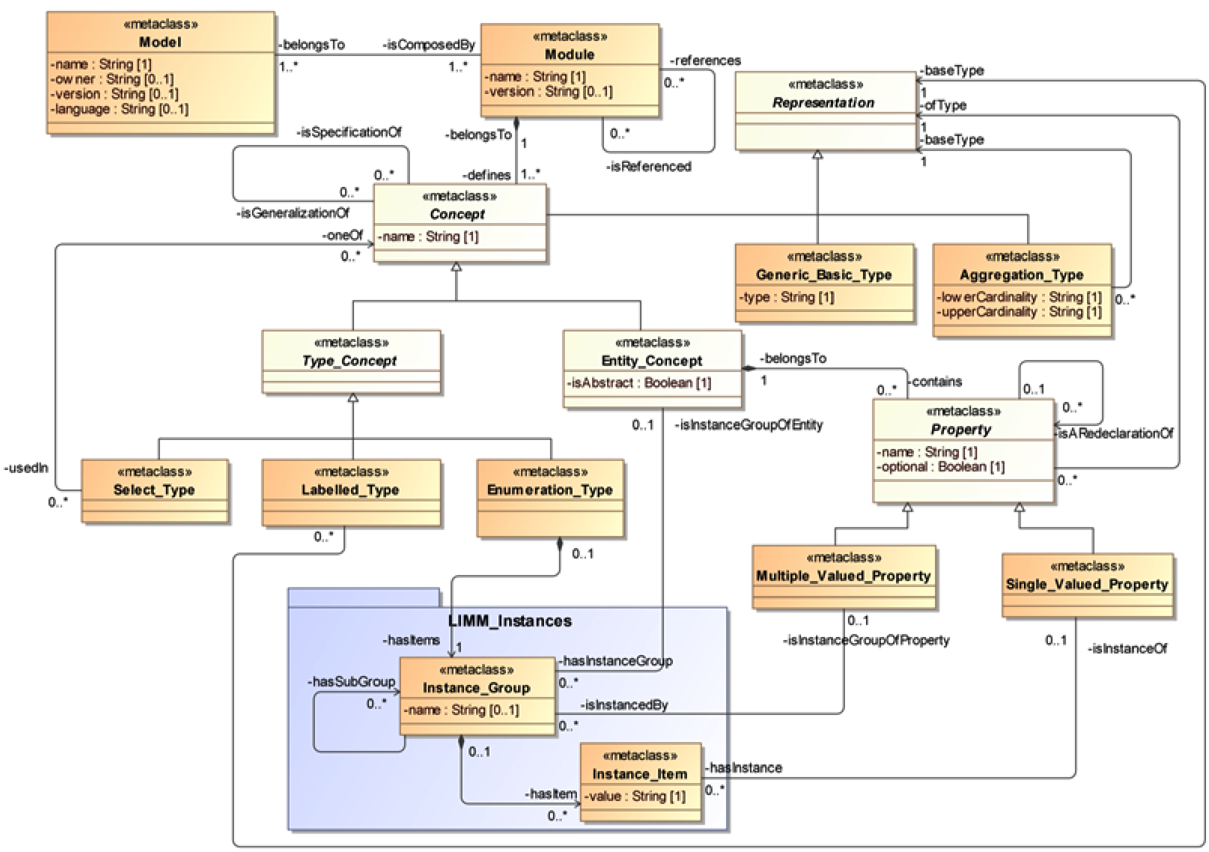


Figure 10.1 - Language Independent Meta-model (LIMM)

Many of the information modelling languages, e.g. EXPRESS (ISO TC184/SC4 2004), UML class infrastructure (OMG 2011b), OWL and XSD specification (W3C 2009; W3C 2001) have been analysed in detail and they were the focus of the attention to create this comprehensive meta-model and, as far the mappings defined for those languages demonstrate, LIMM is able to support them with little loss of expressiveness. In resemblance to what happens in the OWL language, LIMM is capable of representing both models and data levels of MDA (Level 1 and Level 0, respectively), enabling the combined transformation of both levels at the same time, or each independently if required. With this, not only the meta-model is prepared to deal with harmonization of modelling languages, but is also capable of representing instances of models, meaning that it can be used as an intermediate platform for data harmonization (represented by the “LIMM_Instances” package depicted in Figure 4.15 on the bottom).

Concerning modelling concepts, the meta-model considers the representation of entities, types, properties, basic types, aggregations, etc. Nevertheless, some languages (e.g. EXPRESS) enable explicit behavioural expressions (instantiation rules) and functions, which are not supported. However, they are considered non-fundamental for the envisaged mapping process which is mainly focused on the information model mapping at the Level 1 of the framework. A more detailed explanation of the composition of LIMM is presented, evidencing the use of each structure defined:

- <u>*Model*</u>**:** identifies the “header” of the original model, in terms of owner, version and original modelling language. A *Model* can be composed by a multitude of *Modules*;
- <u>*Module*</u>**:** each *Module* represents a fraction or the whole model, since original models can be distributed by a series of resources. The *Module* class identifies by name and version each part of the original model and is composed by *Concepts*;
- <u>*Concept*</u>**:** is an abstract class to represent any kind of modelling structure defined at the root of the module (root elements of the original model representation). It either can be instantiated as complex entities (*Entity_Concept*) or as type declarations (*Type_Concept*);
- <u>*Entity_Concept*</u>**:** this class represents an important structural part of the information model defining classes of objects. An *Entity* contains *Properties* and, when not abstract, can be used to represent palpable real life information, thus it can also be instantiated with level 0 data (through *Instance_Group*). Being a class that enables the specification of other classes, detailing a hierarchy of *Entities*, it allows to mark them as abstract if they are not meant to be directly instantiated;
- <u>*Property*</u>**:** acts as complementary information about a given *Entity_Concept*, since a *Property* cannot exist without a belonging *Entity*. *Properties* have a given underlying associated type which can be any class inherited from the abstract *Representation* class. Moreover, depending of this type, i.e. depending whether it can assume a single value (e.g. an integer) or an aggregation of values (e.g. an array on integers), a *Property* is concretised as *Single_Valued_Property* or *Multiple_Valued_Property*. Redeclared *Properties* should be used only in case of need to redefine some other *Property* already specified within a supertype *Entity_Concept*. With this, a particular *Property* can be renamed and / or type redefined / refined. Finally, a *Property* can be linked with level 0 instances (through

Single_Valued_Property or *Multiple_Valued_Property*) to represent palpable real life information.

- *Single_Valued_Property*: is a specific *Property* that can have at maximum a single instance value. It is connected to *Instance_Item*.
- *Multiple_Valued_Property*: is a specific *Property* that can capture multiple instance values at the same time, as is the case of arrays or matrices. It is connected to *Instance_Group*.
- *Instance_Item*: is the valued instance of a real life concept;
- *Instance_Group*: acting as an aggregator of *Instance_Items*, the *Instance_Group* class represents disjointedly either an instance of an *Entity_Concept*, an aggregation of the possible values of an *Enumeration_Type* or even the values of a property whose underlying type is an *Aggregation_Type*;
- *Representation*: is the top abstract class which can go from *Generic_Basic_Type*, advanced *Types*, passing through *Entities* and *Aggregations*. This class represents the top level of abstraction of a single piece of information that can be modelled by the meta-model or that exists natively in modelling languages (e.g. "Strings");
- *Generic_Basic_Type*: this class represents a predefined language type. It can represent any basic type, which the model demands (e.g. "Integer", "String", "Boolean", etc.);
- *Aggregation_Type*: like the name explains, is a class to represent *Aggregations* (i.e. arrays, bags, vectors, etc.), which can be limited by the "upperCardinality" and "lowerCardinality". This class has no direct information about possible contents besides the type it is associated with. This means that there is no information about possible order and duplicity of elements;
- *Type_Concept*: is an abstract class and represents the high level abstraction of selectors, renamed concepts and enumerations, such as *Select_Type*, *Labelled_Type* and *Enumeration_Type*. Together with the *Entity_Concept*, *Type_Concept* acts as a declaration of model structures, and helps the understanding of the difference which is inherent between them since it cannot have any inner elements such as *Properties*;
- *Select_Type*: is one of the advanced type structures and allows a given property to assume a multitude of different types, not limiting the instantiation to one particular type. This notion of selection only exists on the EXPRESS modelling language;
- *Labelled_Type*: allows to rename any previously defined concept or a native *Representation*;
- *Enumeration_Type*: the last advanced type, defines the use of *Enumerations*, which by definition are a set of well defined named values. These values are inherently constant and, in the LIMM, each *Enumeration* value is considered to be an *Instance_Item* attached to a specific *Instance_Group* representing the scope of all values this type can represent.

To better understand the inwards of how a normal information model looks like as a LIMM, Figure 10.2 illustrates a simple example of a model to represent geometry, more specifically bi-dimensional triangles which are to be specified by a set of 3 points. Shortly and not exhaustively: the example *Model* is composed by a single *Module* which defines 4 *Entity_Concepts* and 1 *Enumeration_Type*; the entity to represent the triangle object is in fact a specification of the abstract concept named "GeometricShape" (which in a larger example can be extended by more *Entity_Concepts*, e.g. line, square, etc.), thus besides the *Multiple_Valued_Property* defined to capture the 3 vertices values, it also contains an indication of the type ("ofType") of the geometric shape given by the *Enumeration_Type* defined; and finally, the entity to represent a bi-dimensional point object defines a *Single_Valued_Property* for both the "x" and "y" coordinates, as well as another one to assign a label to a future instance of the point.

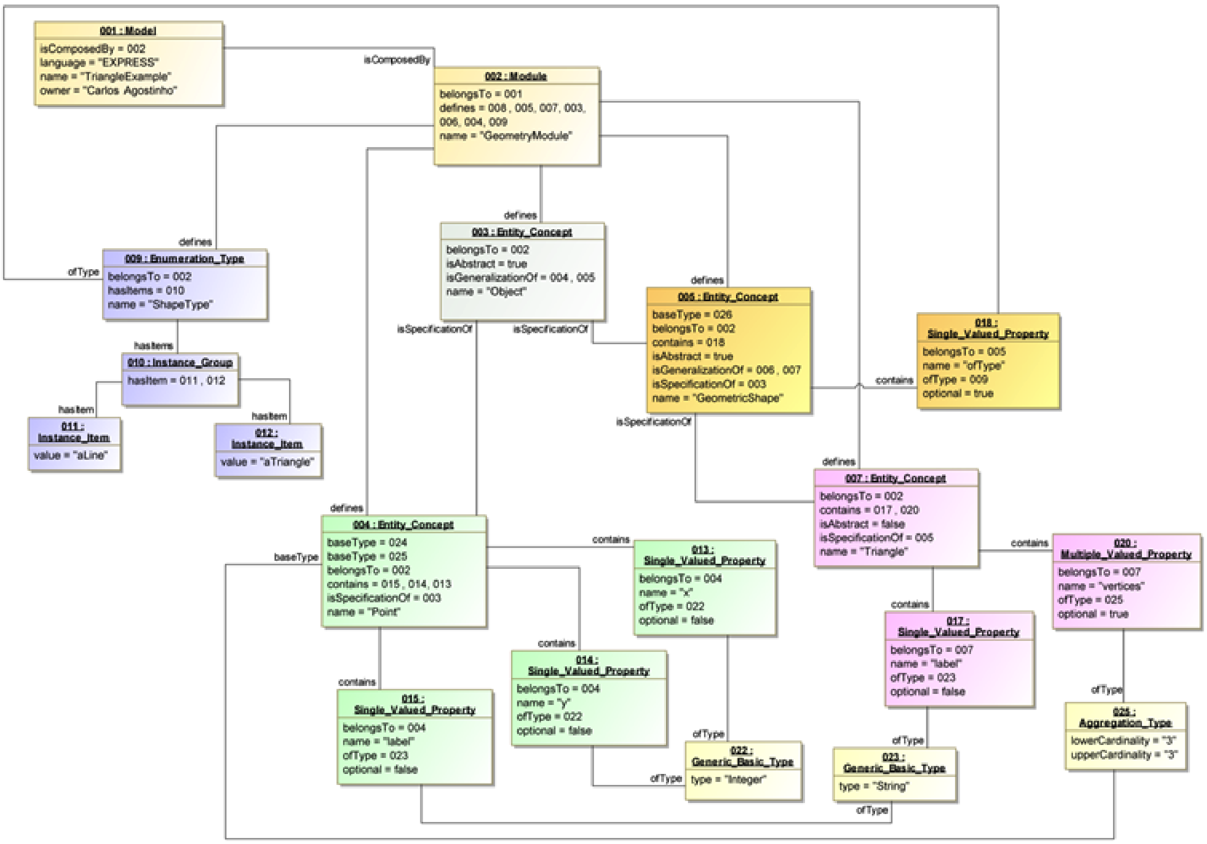


Figure 10.2 – Example of a Simple Language Independent Model

10.2 Course Synopsis Excerpt Example

CRESCENDO	Training Course Synopsis
20 October 2010	Contact: Christian Johansson, TL1.5.2, christian.johansson@ltu.se Tel: +46 (0)70 – 934 38 73 João Sarraipa, TL1.5.3, jfss@uninova.pt Tel: +351 21 2948529

The Crescendo WP1.5 has the objective of developing and organise relevant BDA training activities, targeted towards "future users" (in a broad sense). The BDA training courses and material are intended to help a range of users to develop competence in how to use the results (new knowledge) generated by the Crescendo technical work packages. Hence the training courses will support the consortium partners in the deployment and exploitation of Crescendo results after the end of the project.

This document goal is to describe a training course and its learning objectives. The learning objectives represent the essential understanding that a trainee will reach from attending such course. This document has to be able to answer the question:

What should students know, understand, and be able to do?

The answer to this question will specify what new roles and competencies the trainees will get from attending such course. The course synopsis defined will be essential to conduct training developers to frame and guide the course contents composition to the overall identified course's objective. The synopsis is organised in to two views: Courses and Modules. It follows the idea that a Training Course could be composed by small training objects – the Training Modules.

Course's View Synopsis

Title *[provide a concise title of the proposed **Training Course (TC)**]*:

Systems Engineering Database (SEDb)

Narrative summary *[presents a summary of the TC, its objective and highlights]*:

This course presents the System Architecture, the basic operation principles and the dashboard concept of the CRESCENDO Systems Engineering Database (SEDb).

It addresses the skills about the SEDb:
How to get started (installation guidelines);
How to use the viewer;
How to use the viewer for reporting;
How to use DOORS Database.

Keywords *[presents a set of keywords that could represent the topics addressed by the courses – the main purpose is to use keywords from the CRESCENDO Glossary]*:

System Engineering Database; System Architecture; BDA;

Target Audience *[Response could be one or several of the following categories]*

☐ General Public

☒ Aero industry

☒ IT Vendor people developing, supporting and training on tools to implement BDA Capability

☐ Academia wanting to learn about BDA methods and technologies

Roles & Competences *[Competences to which this course is related with - could be one or several of the following categories]*:

☐ A/C Program Architects & Chief Engineers *[Responsible for setting, and monitoring, the overall objectives for a project/study. Includes Chief Engineers' team (e.g. creation of multiple trade studies from a single objective)]*

☐ A/C program Collaboration Manager *[Responsible for the delegation, and monitoring, of activities throughout an enterprise (both internal and external)]*

☒ Engineering Processes, Method & Tool capability owners *[Responsible for creating and updating methods and tools used within their domain of expertise. Includes configuration & requirements management]*

☒ Design, Modelling & Simulation Engineers *[Responsible for performing the specific methods and tools required to undertake a modelling and simulation activity]*

☐ Quality Auditors *[Responsible for undertaking quality activities required]*

☐ "BDA" Administrators *[Responsible for the administration of the BDA components]*

☐ N/A *[Not applicable to none of the presented Roles]*

Recommended Precedence *[Any CRESCENDO course which the trainee should follow before attending this one]*:

None

Student requirements *[e.g. previous students' knowledge not present in other CRESCENDO course]*:

None

Technical requirements *[e.g. specific installed software]*:

For installation of the SE database following System Requirements should be considered:

System: Windows XP/Vista; RAM: 512MB; Free disk space: 100MB

Estimated Time *[Duration of this TC in minutes]*:

1 hour and 35 minutes

Input Content *[input content sources used as relevant and essential for development of the training contents]*:

D1.2.2

Training Modules: *[This presents the training modules which this TC is composed by]*

Module 0 - Introduction to the Systems Engineering Database (SEDb)

Module 1 - How to get started with the SEDb Viewer

Module 2 - How to use the Viewer

Module 3 – Uncertainty Management

Module 4 - Conclusions

Skills – Know *[Skills to be acquired related to the knowing and understanding (Theoretical knowledge of a field; the capacity to know and understand)]*:

To know the SEDb system architecture.

Skills – Do *[Skills to be acquired related to the knowing how to act (Practical and operational application of knowledge to certain situations; be able to accomplish)]*:

Be able to Install SEDb; to use SEDb viewer; to use the viewer for reporting; and to use DOORS Database; to perform uncertainty management.

Skills – Be *[Skills to be acquired related to the knowing how to be (Values such as an integral element of the way of perceiving and living with others and in a social context)]*:

N/A

Contact Person *[who is the main responsible/author of this TC]*:

N/D

.Modules' View Synopsis

Module 0

Title *[provide a concise title of the **Training Module (TM)**]*:

Introduction to the Systems Engineering Database (SEDb)

Narrative summary *[presents a summary of the TM, its objective and highlights]*:

This module provides a brief introduction to the Systems Engineering Database, including a description of the SEDb's Architecture and Data Model.

Keywords *[presents a set of keywords that could represent the topics addressed by this TM – the main purpose is to use keywords from the CRESCENDO Glossary]*:

SEDb Architecture; SEDb Data Model

Target Audience *[Response could be one or several of the following categories]*

☐ General Public

☒ Aero industry

☒ IT Vendor people developing, supporting and training on tools to implement BDA Capability

☐ Academia wanting to learn about BDA methods and technologies

Roles & Competences *[Competences to which this TM is related with - could be one or several of the following categories]*:

☐ A/C Program Architects & Chief Engineers *[Responsible for setting, and monitoring, the overall objectives for a project/study. Includes Chief Engineers' team (e.g. creation of multiple trade studies from a single objective)]*

☐ A/C program Collaboration Manager *[Responsible for the delegation, and monitoring, of activities throughout an enterprise (both internal and external)]*

☒ Engineering Processes, Method & Tool capability owners *[Responsible for creating and updating methods and tools used within their domain of expertise. Includes configuration & requirements management]*

☒ Design, Modelling & Simulation Engineers *[Responsible for performing the specific methods and tools required to undertake a modelling and simulation activity]*

☐ Quality Auditors *[Responsible for undertaking quality activities required]*

☐ "BDA" Administrators *[Responsible for the administration of the BDA components]*

☐ N/A *[Not applicable to none of the presented Roles]*

Recommended Precedence *[Any CRESCENDO TM which the trainee should follow before attending this one]*:

None

Student requirements *[e.g. previous students' knowledge not present in other CRESCENDO TC or TM]*:

None

Technical requirements *[e.g. specific installed software]*:

None

Estimated Time *[Duration of this TM in minutes]*:

10 minutes

Input Content *[input content sources used as relevant and essential for development of the training contents]*:

D1.1.2

Skills – Know *[Skills to be acquired related to the knowing and understanding (Theoretical knowledge of a field; the capacity to know and understand)]*:

To know the SEDb system architecture and SEDb Data Model

Skills – Do *[Skills to be acquired related to the knowing how to act (Practical and operational application of knowledge to certain situations; be able to accomplish)]*:

N/A

Skills – Be *[Skills to be acquired related to the knowing how to be (Values such as an integral element of the way of perceiving and living with others and in a social context)]*:

N/A

Contact Person *[who is the main responsible/author of this TM]*:

N/D

10.3 Author's Publications List

10.3.1 **Publications in Journals**

1. Jardim-Gonçalves, R.; Sarraipa, J.; and Agostinho, C.; and Panetto, H. (2009). "Knowledge Framework for Intelligent Manufacturing Systems". In: Journal of Intelligent Manufacturing, October 2009, Vol. 20(5), ISSN: 1572-8145.

2. Sarraipa, J.; Jardim-Goncalves, R.; and Steiger-Garcao, A. (2010). "Mentor: an Enabler for Interoperable Intelligent Systems". In: International Journal of General Systems, 39: 5, 557 — 573, First Published On: 13 May 2010.

3. Jardim-Goncalves, R.; Agostinho, C.; Sarraipa, J.; Grilo, A.; and Mendonça, J. (2012). "Reference Framework for Enhanced Interoperable Collaborative Networks in Modern Furniture Industrial Organizations: The FunStep Case Study". In: International Journal of Computer Integrated Manufacturing (2012).

10.3.2 **Book Chapters**

1. Jardim-Goncalves, R.; Agostinho, C.; Sarraipa, J.; Togores, A.; Nuñez, M.; and Panetto, H. (2011). Chapter: "Standards Framework for Intelligent Manufacturing Systems Supply Chain". In the Book: Supply Chain Management - New Perspectives, Isbn 978-953-307-633-1, Edited By Sanda Renko.

10.3.3 **Publications in Conferences:**

1. Jardim-Goncalves, R.; Olavo, R.; Sarraipa, J.; and Steiger-Garcao, A. (2004). "Ontology-Based Framework for Enhanced Interoperability in Networked Industrial Environments". In: INCOM2004, 11th IFAC Symposium on Information Control Problems In Manufacturing, Brazil, 2004.

2. Sarraipa, J.; Zouggar, N.; Chen, D.; Jardim-Goncalves, R. (2007). "Annotation for Enterprise Information Management Traceability". In Proceedings of IDETC/CIE ASME 2007.

3. Agostinho C.; Sarraipa J.; D'antonio, F. Jardim-Goncalves, R. (2007). "Enhancing Step-Based Interoperabity Using Model Morphisms". In: 3rd International Conference On Interoperability For Enterprise Software And Applications (I-ESA 2007)

4. Sarraipa, J.; Silva, J. P.; Jardim-Goncalves, R.; Monteiro, A. (2008). MENTOR – Methodology for Enterprise Reference Ontology Development, In: Intelligent Systems, 2008 (Is '08), 4th International IEEE Conference, Pp 6.32-6.40, 2008.

5. Sarraipa, J.; Onofre, S.; Maló, P. and Jardim-Gonçalves, R. (2008). "Inter-Enterprise Collaboration Throughout Ontological Orchestration". In: Echallenges E-2008, 22-24 Oct 2008, Stockholm, Sweden.

6. Maló, P.; Sarraipa, J.; Jardim-Gonçalves, R.; and Steiger-Garção, A. (2008). "A Training Curriculum in Collaboration for Engineering Management". In: IEMC-Europe 2008 International

Engineering Management Conference, 28-30 Jun 2008, Estoril, Portugal.

7. Maló, P.; Sarraipa, J.; Jardim-Gonçalves, R. And Steiger-Garção, A. (2008). “The Cospaces Training System”. In: ICE2008 - 14th International Conference on Concurrent Enterprising, 23-25 Jun 2008, Lisbon, Portugal.

8. Sarraipa, J.; Agostinho, C.; Panetto, H.; and Jardim-Gonçalves, R. (2009). “Semantic Enrichment of Standard-Based Electronic Catalogues”. In: 13th IFAC Symposium on Information Control Problems in Manufacturing (IFAC-INCOM), Jun 2009, Moscow, Russia.

9. Silva, J.; Monteiro, A.; Sarraipa, J.; and Jardim-Goncalves, R. (2009). “Towards Ontology Harmonization Of Mechanical Manufacture Constraints Through PLC”. In: Proceedings of The 20th International Congress Of Mechanical Engineering, November, 2009, Gramado-Rs Brazil.

10. Sarraipa, J.; Vieira, H.; Agostinho, C.; Goncalves, R. (2009). “Methodology for Seamless Supply Chain Planning”. In: 1st International Conference on Software, Services & Semantic Technologies (S3t), 28-29 Oct 2009, Sofia, Bulgaria.

11. Sarraipa, J.; Gaspar, T.; Jardim-Goncalves, R.; and Steiger-Garcao, A. (2010). “Collaborative Ontology Building Using Qualitative Information Collection Methods”. In: 5th IEEE International Conference on Intelligent Systems IS’10, July 7-9, London, Uk.

12. Sarraipa, J.; Figueiredo, D.; Maló, P.; and Jardim-Goncalves, R. (2010). “An Inter-Organisational Approach to Industrial E-Training”. In: Collaborative Knowledge Management - CKM 2010: Inter-Organizational Collaboration, Collaborative Software Development, And Knowledge Management (I) - Invited Session (0000014) (KGCM) Of The 3rd International Multi-Conference On Engineering And Technological Innovation: IMETI 2010.

13. Jardim-Goncalves, R.; Sarraipa, J.; and Steiger-Garcao, A. (2010). “Semantic Harmonization for Seamless Networked SCP in the Future of Internet”. In: Enterprise Architecture, Integration, Interoperability and Networking (EAI2N) Of the 17 Ifip International Conferences held at September 20-23 2010 In Brisbane Australia.

14. Jardim-Goncalves, R.; Sarraipa, J.; And Silva, J. (2010). Methodology for The Development Of Sustainable Reference Ontology In The Mechanical Engineering Domain. In Proceedings of the ASME 2010 International Mechanical Engineering Congress & Exposition - IMECE2010, November 12-18, 2010, Vancouver, British Columbia, Canada

15. Jardim-Goncalves, R.; Sarraipa, J.; And Silva, J. (2010). “Sustainable Reference Ontology Development in Mechanical Engineering Domain”. In Proceedings of the ASME 2010 International Mechanical Engineering Congress & Exposition - Imece2010, November 12-18, 2010, Vancouver, British Columbia, Canada.

16. Agostinho, C.; Sarraipa, J., Goncalves, D., And Jardim-Goncalves, R. (2011). “Tuple-Based Semantic and Structural Mapping for a Sustainable Interoperability”. In: DOCEIS’11 2nd Doctoral

Conference on Computing, Electrical and Industrial Systems – Costa De Caparica, Lisbon, February 2011.

17. Sarraipa, J.; Jardim-Goncalves,R. (2011). “Knowledge-Based System for Semantics Adaptability of Enterprises Information Systems”. In: IWEI 2011 - The International IFIP Working Conference on Enterprise Interoperability. March 23-24, on Stockholm, Sweden.

18. Sarraipa, J.; Jardim-Goncalves, R. (2011). “Semantics Adaptability for Systems Interoperability”. In Proceedings of the 1st UNITE Doctoral Symposium, 27-28 June 2011, Bucharest, Romania

19. Beca, M.; Sarraipa, J.; Agostinho, C.; And Jardim-Goncalves, R. (2011). “Tuple-Based Morphisms for e-Procurement Solutions”. In: Proceedings of the ASME 2011 International Design Engineering Technical Conference & Computers And Information In Engineering Conference IDETC/CIE 2011 (2011)

20. Ferreira, J.; Agostinho, C.; Sarraipa, J., And Jardim-Goncalves, R. (2011). “Monitoring Morphisms to Support Sustainable Interoperability of Enterprise Systems”. In: Proceedings of the EI2N 2011 : 6th International Workshop On Enterprise Integration, Interoperability and Networking, October 19-20 2011 Crete, Greece

21. Silva, J.; Cavaco, F.; Sarraipa, J.; And Jardim-Goncalves, R. (2011). “Knowledge Based Methodology Supporting Interoperability Increase in Manufacture Domain”. In the Proceedings of the ASME Congress2011 ASME International Mechanical Engineering Congress and Exposition IMECE 2011, November 11-17, 2011, Denver, CO, Usa

22. Ferreira, J.; Agostinho, C.; Sarraipa, J., and Jardim-Goncalves, R. (2012). “Monitoring Morphisms to Support Sustainability of Interoperability in the Manufacturing Domain”. In the Proceedings of the INCOM 2012 - 14th Ifac Symposium on Information Control Problems in Manufacturing, Bucharest, Romania, May 23-25, 2012.

23. Restrepo, E. G.; Medina, R.; Briones, R.; Córdova, I.; Medina, G.; Pinzón, S.; Coronado, P.; Finat, C.; Bticario, J. G.; Méndez, D.; Sacco, A.; Vanegas, C.; Andrade, F.; Rodríguez, S.; Ginocchio, M. I. ; Membreño, O. Z.; Sarraipa, J. (2012). “Kit Alter-Nativa: Empoderando A Los Profesores Para Una Educación En Contextos De Diversidad”. In: The III Congreso Iberoamericano Sobre Calidad Y Accesibilidad De La Formación Virtual (CAFVIR 2012), Alcalá De Henares, España, April 25-27, 2012.

24. Sarraipa, J.; Baldiris, S., Fabregat, R.; And Jardim-Goncalves, R. (2012). “Knowledge Representation in Support of Adaptable Elearning Services for All. In the Proceedings of the DSAI’12 - Software Development for Enhancing Accessibility and Fighting Info-Exclusion, Douro Region, Portugal, 19-22 July, 2012.

25. Matzakou, I.; Askounis, D.; Sarraipa, J., And Jardim-Goncalves, R. (2012). “Enhanced Knowledge Management of an Interoperability Service Utility”. In the proceedings of the 6th IEEE International Conference On Intelligent Systems IS’12, Sofia, Bulgaria, September 6-8, 2012.

26. Alves, G., Sarraipa J., Silva, J. P. M and Jardim-Gonçalves R. A (2012). “Framework for Semantic Checking of Information Systems”. In the: Fifth Interop-Vlab.It Workshop, Co-Located With Itais 2012 (IX Conference of the Italian chapter of AIS Organization Change And Information Systems: Working and Living Together in New Ways), 28th Of September 2012 In Rome, Italy (2012).

27. Marcelino-Jesus, E.; Sarraipa, J.; Antão, M.; Jardim-Gonçalves R.; (2012). “Methodology for the Economic Viability of Companies in the Semantic Adaptation of Information Systems”. In the: ASME International Mechanical Engineering Congress And Exposition, IMECE 2012, November 9-15, 2012, Houston, TX, Usa.

11 References

[1] Jardim-Goncalves, R., Grilo, A., & Steiger, A. (2006). “Challenging the interoperability between computers in industry with MDA and SOA”. In: Computers in Industry, 57(8–9), December 2006, 679–689.

[2] Panetto, H., Jardim-Gonçalves, R., & Pereira, C. (2006). “EManufacturing and web-based technology for intelligent manufacturing and networked enterprise”. In: Journal of Intelligent Manufacturing,17(6), 639–640.

[3] Online Etymology Dictionary (2010). “Inter-“ meaning, retrieved from the web: http://www.etymonline.com/index.php?search=inter&searchmode=none, at August 2010.

[4] Your Dictionary.com (2010). “Inter“ meaning, retrieved from the web: http://www.yourdictionary.com/computer/inter, at August 2010.

[5] Online Etymology Dictionary (2010). “-ability“ meaning, retrieved from the web: http://www.etymonline.com/index.php?search=-ability&searchmode=none, at August 2010.

[6] The Free Dictionary (2012). “Operation” meaning, retrieved from the web at April 2012: http://www.thefreedictionary.com/operation.

[7] IEEE Standard Computer Dictionary (1990). “Interoperability”. In: A Compilation of IEEE Standard Computer Glossaries; New York, NY: 1990.

[8] Chen, D.; Doumeingts, G.; Vernadat, F.; (2008). “Architectures for enterprise integration and interoperability: Past, present and future”, in: Computers in Industry, v.59 n.7, p.647-659, September, 2008.

[9] IDEAS Consortium (2003). IDEAS Project Deliverables (WP1-WP7), Public Reports, 2003, www.ideasroadmap.net.

[10] Panetto, H. (2007). “Towards a classification framework for interoperability of enterprise applications”. In: International Journal of Computer Integrated Manufacturing, Vol. 20, No. 8, December 2007, 727 – 740.

[11] Chen, D.; Doumeingts, G. (2003). “European initiatives to develop interoperability of enterprise applications—basic concepts, framework and roadmap”, In Journal of Annual reviews *in Control* 27 (3) (2003), pp. 151–160.

[12] The Free Dictionary (2010). Framework Meaning. Retrieved from the web at August 2010: http://www.thefreedictionary.com/framework.

[13] Work Package – A4.2 Participants (2007). Athena Deliverable Number: “D.A4.2: Specification of Interoperability Framework and Profiles, Guidelines and Best Practices – version 1.0”; March, 2007.

[14] European Communities (2004). “European Interoperability Framework for Pan-European eGovernment Services”; IDA working document – Version 4.2; 2004.

[15] INTEROP Partners (2003). Annex I - “Description of Work”; November 2003.

[16] Li, Man-Sze; Cabral, R.; Doumeingts, G.; Popplewell, K. (2006). “Enterprise Interoperability Research Roadmap”, European Commission, Final Version (Version 4.0) 31 July 2006.

[17] Charalabidis, Y; Gionis, G; Hermann, K; Martinez, K. (2008). “Enterprise Interoperability: Research Roadmap”. Update Version 5.0. March 5th, 2008.

[18] Commission of the European Communities (2005). Communication from the Commission to the Council, The European Parliament, The European Economic and Social Committee and the Committee of the Regions “i2010 – A European Information Society for growth and employment” - {SEC(2005) 717}; Brussels, 1.6.2005.

[19] ISO. (2002). ISO 16100-1:2002. Industrial automation systems and integration - Manufacturing software capability profiling for interoperability - Part 1: Framework.

[20] International Standard ISO 1087 (1990). Terminology – Vocabulary, 1990.

[21] INTEROP NoE Project members (2006). Semantic Interoperability definition – Interoperability ontology – INTEROP NoE Project (2006).

[22] Jardim-Goncalves, R.; Sarraipa, J.; Agostinho, C.; Panetto, H. (2009). “Knowledge Framework for Intelligent Manufacturing Systems”. In: International Journal of Intelligent Manufacturing, 2009

[23] Venkataram, P. (2009?). In the article: “An Article to clear up some Misconceptions about the nature of Research”; retrieved from the web: http://cce.iisc.ernet.in/motivationprinciples.pdf at August, 2010.

[24] Camarinha-Matos, L. (2009). “Unit 2: Scientific Method”. In slides of the Scientific Research Methodologies and Techniques course of the PhD Program in Electrical and Computer Engineering of the FCT-UNL - 2009.

[25] Meystel, A.M. and Albus, J.S., (2001). “Intelligent systems: architecture, design, and control”. New York: John Wiley and Sons.

[26] Saridis, G.N. and Valavanis, K.P., (1988). “Analytical design of intelligent machines”. Automatica, 24 (2), 123–133.

[27] Zadeh, L.A. (1994). “Fuzzy logic, neural networks, and soft computing”. Communications to the ACM, 37 (3), 77–84. ISSN:0001-0782.
[28] Kasabov, N.; and Filev, D. (2006). “Evolving intelligent systems: methods, learning, & applications”. In Proc. Int. Symposium on Evolving Fuzzy Systems, pp. 8–18.
[29] Sarraipa, J.; Jardim-Goncalves, R. and Steiger-Garcao, A. (2010). “MENTOR: an enabler for interoperable intelligent systems”, International Journal of General Systems, 39: 5, 557 — 573, First published on: 13 May 2010 (iFirst)
[30] Gruber, T. (1993). “Toward Principles for the Design of Ontologies Used for Knowledge Sharing”. Originally in N. Guarino and R. Poli, (Eds.), International Workshop on Formal Ontology, Padova, Italy. Revised August 1993. Published in International Journal of Human-Computer Studies, Volume 43, Issue 5-6 Nov/Dec. 1995, pp 907-928, (1993).
[31] Hayek, F. (1945). “The use of knowledge in society”. The American economic review, 35 (4), 519–530.
[32] May, M. (2003). “Business Process Management: Strategic Integration in a web-enabled environment”. Pearson Education. Ebook: Financial Times Management. ISBN: 0-273-66108-6.
[33] Maló, P.; Sarraipa, J.; Jardim-Gonçalves, R.; and Steiger-Garção, A. (2008). “A Training Curriculum in Collaboration for Engineering Management”. In: IEMC-Europe 2008 International Engineering Management Conference, 28-30 Jun 2008, Estoril, Portugal.
[34] Hammond, J., Koubek, R.J., Harvey, C.M. (2001). “Distributed collaboration for engineering design: A review and reappraisal”. International Journal of Human Factors in Manufacturing, 11, 35-52.
[35] Toikka, S. (2007). “ICT Supported Interorganizational Knowledge creation: Application of Change Laboratory”. In proceedings of ECTEL 2007 - Second European Conference on Technology Enhanced Learning, 17-20 September 2007, Crete, Greece.
[36] Tieju, M. (2007). “Knowledge Creation Models and Knowledge Maps”. Retrieved from the web at May 2011: http://www.jaist.ac.jp/coe/english/project/pdf/04_Ae.pdf.
[37] Nonaka, I., Takeuchi, H. (1995). In the Book: “The Knowledge-Creating Company”. Oxford University Press.
[38] Paavola, S.; Lipponen, L; and Hakkarainen, K. (2002). “Epistemological Foundations for CSCL: A Comparison of Three Models of Innovative Knowledge Communities”. Published in: CSCL '02 Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community.
[39] Bereiter, C. (2002). Book Title: “Education and Mind in the Knowledge Age”. Publisher: Lawrence Erlbaum Associates. Place of Publication: Mahwah, NJ. Publication Year: 2002.
[40] Engeström, Y. (1987). “Learning by expanding”. Orienta-Konsultit Oy, Helsinki.
[41] Engeström, Y. (1999). “Innovative learning in work teams: Analyzing cycles of knowledge creation in practice”. In Y. Engeström, R. Miettinen & R.-L-. Punamäki (Eds.), Perspectives on activity theory, (pp. 377-404). Cambridge: Cambridge University Press.
[42] The OR Society (2011). Staff turnover is an increasing problem. Are you allowing your company's valuable knowledge assets to walk out the door? Retrieved from the web at May 2011: http://www.orsoc.org.uk/learn/IssueExplorer/explor2.htm.
[43] Life Sciences Consulting Partners (2011). Internalizing Enterprise Knowledge Assets in your decision process. Retrieved from the web at May 2011: http://www.lsconsultingpartners.com/index.php?option=com_content&view=article&id=75&Itemid=64.
[44] KADI - Knowledge Assets Development, Inc. (2011). “Improving Performance via What-If Process Planning”. Retrieved from the web at May 2011: http://www.knowledgeassets.com/.
[45] Wikipedia (2010). Knowledge meaning. Retrieved from the web: http://en.wikipedia.org/wiki/Knowledge, at August 2010.
[46] Wikipedia (2010). Plato description. Retrieved from the web: http://en.wikipedia.org/wiki/Plato, at August 2010.
[47] Wikipedia (2011). Justified true belief. Retrieved from the web at August 2011: http://en.wikipedia.org/wiki/Justified_true_belief
[48] Bellinger, G., Castro, D., Mills, A. (2004). “Data, Information, Knowledge and Wisdom”. Retrieved from the web http://www.systems-thinking.org/dikw/dikw.htm at August 2010.
[49] Jiang Hua (2008). “Study on Knowledge Acquisition Techniques”. In: 2008 Second International Symposium on Intelligent Information Technology Application
[50] Mueller, H. A. Jahnke, J. H.; Smith, D. B.; Storey M. A. (2000). “Reverse Engineering: A Roadmap”. In: Proc. Int’l Conf. Software Eng. (ICSE-00), ACM Press, New York, 2000, pp. 47–60.
[51] Maedche, Alexander; Staab, Steffen (2001). "Ontology Learning for the Semantic Web". In: IEEE Intelligent Systems, vol. 16, no. 2, pp. 72-79, March/April, 2001.

[52] Dom, B. (1999). “Automatically finding the best pages on the World Wide Web (CLEVER)”. In Search Engines and Beyond: Developing efficient knowledge management systems, April 19–20 1999, Boston, Mass.
[53] Brewster, C.; Ciravegna, F. and Wilks, Y. (2001). “Knowledge Acquisition for Knowledge Management”. Position Paper. In Proceedings IJCAI 2001 Workshop on Ontology Learning, Seattle, USA.
[54] Rhem, A.J. & Associates, Inc. (2001). “A Framework for Knowledge Acquisition”. White Paper May 2001. Retrieved from the web at August 2011: http://www.ajrhem.com/framework.pdf
[55] Navigli, R., & Velardi (2008). “From glossaries to ontologies: Extracting semantic structure from textual definitions”. In Proceedings of the 2008 conference on ontology learning and population: Bridging the gap between text and knowledge (pp. 71–87). Amsterdam, The Netherlands: IOS Press.
[56] Buitelaar, P., Cimiano, P., & Magnini, B. (2005). “Ontology learning from text: Methods, evaluation and applications”. Frontiers in Artificial Intelligence and Applications Series (Vol. 123). IOS Press.
[57] Clark, M.; Kim, Y.; Kruschwitz, U.; Song, D.; Albakour, D.; Dignum, S.; Beresi, U. C.; Fasli, M.; Roeck, A. D. (2011). “Automatically structuring domain knowledge from text: An overview of current research”. In: Information Processing & Management (August 2011).
[58] Sclano, F.; Velardi, P., (2007). “TermExtractor: a Web Application to Learn the Common Terminology of Interest Groups and Research Communities”. In Proc. of 9th Conf. on Terminology and Artificial Intelligence (TIA 2007). Sophia Antinopolis.
[59] Cimiano, P.; Hotho, A.; and Staab, S. (2005). “Learning concept hierarchies from text corpora using formal concept analysis”. In: Journal of Artificial Intelligence Research, 24 (2005), pp. 305-339.
[60] Landauer, T. K. & Dumais, S. T. (1997). “A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge”. Psychological Review, 104, 211-14
[61] Hoffman, R.R.; Shadbolt, N.R.; Burton, A.M.; and Klein, G. (1999). “Eliciting knowledge from experts: a methodological analysis”. Organisational behaviour and human decision processes. Vol. 62. No. 2. pp. 129-158.
[62] Garcia-Perez, A.; and Mitra, A. (2007). “Tacit Knowledge Elicitation and Measurement in Research Organisations: a Methodological Approach”. In: Electronic Journal of Knowledge Management Volume 5 Issue 4 2007 (373-386)
[63] InSites (2007). Tips on Qualitative and Quantitative Data Collection Methods. Retrieved from the web at August 2011: http://www.insites.org/CLIP_v1_site/downloads/PDFs/TipsQualQuanMthds.4B.8-07.pdf
[64] Borysowich, C; (2006). Using the Nominal Group Method to Drive a Team Decision. Retrieved from the web, at March 2010: http://it.toolbox.com/blogs/enterprise-solutions/using-the-nominal-group-method-to-drive-a-team-decision-11774
[65] Sample, J. A.; (1984). “Nominal Group Technique: An Alternative to Brainstorming”. In: Journal of Extension, Volume 22, Number 2, ISSN 1077-5315, March 1984.
[66] JKU- INTERNET (2010). “Metaplan”. Retrieved from the web at March 2010: http://elearn.jku.at/wiki/index.php/Metaplan.
[67] McBurney, D. H.; White, T. L.; (2007). “Nonexperimental Research, Part 2: Survey Research”, in Resarch Methods (8th edition), Chapter 10, pp. 245-246, 2007.
[68] Linstone, H.; Turoff, M.; (1975). The Delphi Method: Techniques and Applications, Addison-Wesley, Reading, MA, 1975
[69] TechPluto (2011). Core Characteristics of Web 2.0 Services. In the web, accessed at August 2011: http://www.techpluto.com/web-20-services/.
[70] Wikipedia (2011). WEB 2.0 definition. Retrieved from the web at August 2011: http://en.wikipedia.org/wiki/Web_2.0
[71] Kaminski, J. (2009). “Folksonomies boost Web 2.0 Functionality”. In: OJNI Online Journal of Nursing Informatics, 13 (2), Summer 2009
[72] Halpin, H.; Robu, V. Shepherd, H. (2007). “The complex dynamics of collaborative tagging”. Proceedings of the 16th international conference on World Wide Web, May 08-12, 2007, Banff, Alberta, Canada.
[73] Sinclair, J. and Cardew-Hall, M. (2008). “The folksonomy tag cloud: when is it useful?” In: Journal of Information Science, 34(1):15–29, 2008
[74] Kyrnin, J. (2011). Tagging: What is Tagging and Why Should We Do It? Learn how to add small data chunks to your Web pages. In the web site About.com. Retrieved at August 2011: http://webdesign.about.com/od/tagging/a/aa020507.htm
[75] Jacob, E. (2004). “Classification and categorization: A difference that makes a difference”. Library Trends, 52(3):515–540, 2004.

[76] OCLC Online Computer Library Center, Inc. (2011). Tag Cloud Builder. Accessed at August 2011: http://tagcloud.oclc.org/tagcloud/TagCloudDemo
[77] Mika, P. (2005). "Ontologies Are Us: A Unified Model of Social Networks and Semantics". In the proceedings of the conference ISWC 2005 – 4th International Semantic Web Conference, Galway Ireland.
[78] Makhfi, P. (2011). Introduction to Knowledge Modeling. Retrieved from the web at August 2011: http://www.makhfi.com/KCM_intro.htm
[79] Auer, S.; Dietzold, S.; Riechert, T. (2006). "OntoWiki – A Tool for Social, Semantic Collaboration". In Proceedings of the 5th international conference on The Semantic Web (ISWC'06). ISBN: 978-3-540-49029-6, pp 736 - 749, (2006).
[80] The Free Dictionary (2011). "Taxis" definition. Retrieved from the web at August 2011. http://www.thefreedictionary.com/taxis
[81] The Free Dictionary (2011). "Folk" definition. Retrieved from the web at August 2011. http://www.thefreedictionary.com/folk
[82] Arms, W. (2000). "Lexicon definition after Digital Libraries", by William Arms, 2000 M.I.T. Press.
[83] Brent, M. R. and Cartwright, T. A. (2002). "Distribution regularity and phonotactic constraints are useful for segmentation". In book: Psycholinguistics: Critical Concepts in Psychology, by Altmann Gerry.
[84] Sarraipa, J.; Silva, J.; Jardim-Goncalves, R. and Monteiro, A. (2008). "MENTOR – A Methodology for Enterprise Reference Ontology Development". In: 2008 4th International IEEE Conference on Intelligent Systems, pp. 6-32 – 6-40.
[85] Wikipedia (2010). "Typology". Retrieved from the web: http://en.wikipedia.org/wiki/Terminology#Typology, at August 2010
[86] CMU (Carnegie Mellon University)-Domain Dictionary. (2007). Accessed 30 Sep 2008. Available on: www.sei.cmu.edu/domainengineering/domain_dict.html.
[87] Wikipedia (2010). "Taxonomy". Retrieved from the web: http://en.wikipedia.org/wiki/Taxonomy, at August 2010.
[88] Malone, J. L. (1988). "The Science of Linguistics in the Art of Translation: Some Tools from Linguistics for the Analysis and Practice of Translation", p. 112 – Published by State University of New york Press, Albany – ISBN 0-88706-653-4.
[89] Soergel (2002). Thesauri in Digital Libraries, JCDL 2002, Portland, OR, 2002-7-14;
[90] Gruber, T.R., (1993). "A translation approach to portable ontology specifications". Journal of knowledge acquisition, 5 (2), 199–220.
[91] Brewster, C.; O'Hara, K.; Fuller, S.; Wilks, Y.; Franconi, E.; Musen, M.; Ellman, J.; Shum, S. (2004). "Knowledge Representation with Ontologies: The Present and Future". In: IEEE Intelligent Systems, vol. 19, no. 1, pp. 72-81, doi:10.1109/MIS.2004.1265889, (2004).
[92] Estival, D.; Nowak, C.; Zschorn, A. (2004). "Toward Ontology-Based Natural Language Processing". In: Proceedings of the Workshop on NLP and XML (NLPXML-2004): RDF/RDFS and OWL in Language Technology, pp 59 - 66, (2004).
[93] Gruber, T. (2005). "Ontology of Folksonomy: A Mash-up of Apples and Oranges". In: Int'l Journal on Semantic Web & Information Systems, 3(2), 2007, (2005).
[94] Saias, J. (2003). "Uma Metodologia para a construção automática de Ontologias e a sua aplicação em Sistemas de Recuperação de Informação" Master thesis available at: http://hdl.handle.net/10174/1560, 2003
[95] Guarino, N.; and Giaretta, P. (1995). "Ontologies and Knowledge Bases". In Proceedings, Towards Very Large Knowledge Bases, Amsterdam, 1995, pp. 25-32.
[96] Daniel E. O'Leary (1998). "Using AI in Knowledge Management: Knowledge Bases and Ontologies". IEEE Intelligent Systems 13 (1998b), pp. 34–39. Retrieved from the web at August 2011: http://lpis.csd.auth.gr/mtpx/km/material/ieee-is-13-3-a.pdf
[97] Sváb, O.; Svátek, V. (2007)."Ontology Mapping enhanced using Bayesian Networks". Retrieved from the web at August 2010: http://nb.vse.cz/~svatek/znal07.pdf
[98] Gómez-Pérez, A.; Fernández-López, M.; Corcho, O. (2002). Technical Roadmap D.1.1.2. OntoWeb Consortium.
[99] Gómez-Pérez, A.; Fernández-López, M.; Corcho, O. (2004). "Ontological Engineering: Methodologies and Tools". The Second European Summer School on Ontological Engineering and the Semantic Web (SSSW-2004).
[100] Pinto, H. S., Gómez-Pérez, A.; Martins, J. P. (1999). "Some Issues on Ontology Integration". In Proceedings of IJCAI99's Workshop on Ontologies and Problem Solving Methods: Lessons Learned and Future Trends, pages 7.1–7.12, 1999.
[101] Missikoff, M.; Boudjlida, N.; Lenzerini, M.; Jeusfeld, M.; Johannesson, P.; et al. (2004). "D8.1 - State of the art and state of the practice including initial possible research orientations", Deliverable D8.1, NoE INTEROP - IST Project n. 508011 - 6th EU Framework Programme (2004).

[102] Lenat, D.B.; Guha, R.V. (1990). “Building Large Knowledge-Based Systems: Representation and Inference in the CYC Project”. Addison Wesley.
[103] Uschold, M.; King, M. (1995). “Towards a methodology for building ontologies”. In proceedings IJCAI-95 Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal, Canada.
[104] Fox, M.S., Gruninger, M., (1997), "On Ontologies and Enterprise Modelling", International Conference on Enterprise Integration Modelling Technology 97.
[105] Bernaras, A.; Laresgoiti, I.; Corera, J. (1996).” Building and Reusing Ontologies for Electrical Network Applications”. Proceedings European Conference on Artificial Intelligence (ECAI’96).
[106] Fernandez, M.; Gomez-Perez, A.; Juristo, N. (1997). “METHONTOLOGY: From Ontological Art Towards Ontological Engineering”. Symposium on Ontological Engineering of AAAI, Stanford, California, 1997, pp. 33-40.
[107] Swartout, R. Patil, K. Knight, T. Russ (1997): “Toward Distributed Use of Large-Scale Ontologies”. Symposium on Ontological Engineering of AAAI, Stanford, California, 1997.
[108] Euzenat, J. (1996). “Corporative memory through cooperative creation of knowledge bases and hyper-documents”. Proceedings 10th KAW, Banff (Canada).
[109] Decker, S.; Erdmann, M.; Fensel, D; Studer, R. (1999). “Ontobroker: Ontology Based Access to Distributed and Semi-Structured Information”. In "Semantic Issues in Multimedia Systems", Proceedings of DS-8, Kluwer Academic Publisher, 1999, pp. 351-369.
[110] Stumme, G.; Madche, A. (2001). “FCA-Merge: Bottom-up merging of ontologies”. In 7th Intl. Conf. on Artificial Intelligence (IJCAI '01), pages 225--230, Seattle, WA, 2001.
[111] Noy, N. F.; Musen, M. A. (2003). “The PROMPT Suite: Interactive Tools for Ontology Merging and Mapping”. International Journal of Human-Computer Studies, Volume 59, Issue 6, December 2003, Pages 983-1024.
[112] Gangemi, A.; Pisanelli, D. M.; Steve, G. (1999). “An Overview of the ONIONS Project: Applying Ontologies to the Integration of Medical Terminologies”. Data & Knowledge Engineering 31, pp. 183-220.
[113] Nicola, A.; Missikof, M.; Navigli, R. (2005). “A proposal for a unified process for ontology building: UPON”. In Proceedings of the 16th international conference on Database and Expert Systems Applications (DEXA), Denmark, LNCS no. 3588, Springer Berlin, pp. 655-664.
[114] NeOn Project, 2010. Acessed at August 2011: http://www.neon-project.org/nw/Welcome_to_the_NeOn_Project
[115] Semantic Web (2010). Retrieved from the web at August 2010: http://semanticweb.org/wiki/OntoStudio
[116] NeOn Toolkit 1.2 Manual (2011). Retrieved from the web at August 2011: http://deki.uca.es/@api/deki/files/324/=NeOn_Toolkit_1_2_Manual.pdf
[117] Stanford Center for Biomedical Informatics Research (2011). Protégé web site. Retrieved from the web at August 2011: http://protege.stanford.edu/
[118] Lin, H. N.; Tseng, S. S.; Weng, J. F.; Lin, H. Y.; Su, J. M.; (2007). “An Iterative, Collaborative Ontology Construction Scheme”. In: Second International Conference on Innovative Computing, Information and Control (ICICIC 2007), ISBN: 0-7695-2882-1, pp. 150, 2007.
[119] Holsapple, C. W.; Joshi, K. D.; (2002). “A Collaborative Approach to Ontology Design”. In: Communications of the ACM, ISSN: 0001-0782, Vol. 45, pp. 42-47, 2002.
[120] Collaborative Protégé Community (2009). Retrieved from the web at Novemeber 2009 : http://protegewiki.stanford.edu/index.php /Collaborative_Protege.
[121] Dietzold, S. (2011). OntoWiki. Retrieved from the wb at August 2011: http://ontowiki.net/Projects/OntoWiki
[122] Poler, R.; Tomás, J.; Velardi, P. et al.; (2004). D10.1 – Interoperability Glossary, Interoperability Research for Networked Enterprises Applications and Software Network of Excellence, 2004.
[123] Borgo, S.; Gangemi, A.; Guarino, N.; Masolo, C.; Oltramari, A.; (2002). “Ontology roadmap”. Retrieved at June 2009 from the web: http://wonderweb.semanticweb.org/deliverables/documents/D15.pdf.
[124] Menzel, C. (2003). “Reference ontologies – application ontologies: either/or or both/and?” Proceedings of the KI2003 workshop on Reference Ontologies and Application Ontologies, 16 September 2003.
[125] Guarino, N.; and Schneider, L., (2002). “Ontology-driven conceptual modeling”, Lecture Notes In Computer Science, Retrived from web: ‘http://www.loa-cnr.it/Tutorials/GuarinoRTutorialPart1.pdf’ at June 2009.
[126] Bechhofer, S.; Gangemi, A.; Guarino, N.; Harmelen, F.; Horrocks, I.; Klein, M.; Masolo, C.; Oberle, D.; Staab, S.; Stuckenschmidt, H.; and Volz, R. (2004). “Tackling the ontology acquisition bottleneck: an experiment in ontology re-engineering”. Available from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.3.2350 [Accessed April 2010].

[127] Agostinho, C.; Jardim-Goncalves, R. (2009). Dynamic Business Networks: A Headache for Sustainable Systems Interoperability. In Proceedings 4th International Workshop on Enterprise Integration, Interoperability and Networking (EI2N'2009) of the On The Move Conference, Vilamoura, Portugal, November (2009)
[128] Lubell, J.; Peak, R. S.; Srinivasan, V.; and Waterbury S. C. (2004), "STEP, XML, AND UML: COMPLEMENTARY TECHNOLOGIES". In: Proc. of DETC 2004, ASME 2004 Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Salt Lake City, Utah USA, Sep 28- Oct 2, 2004.
[129] Agostinho, C.; Delgado, M; Steiger-Garcao, A.; and Jardim-Goncalves, R. (2006). "Enabling Adoption of STEP Standards Through the Use of Popular Technologies". In: 13th ISPE International Conference on Concurrent Engineering (CE 2006), research and applications, 18-22 Sep 2006.
[130] INTEROP NoE (2005). Deliverable MoMo.2 - TG MoMo Roadmap. InterOP, 2005
[131] Agostinho C.; Sarraipa J. D'Antonio F.; Jardim-Goncalves, R. (2007). "Enhancing STEP-based interoperabity using model morphisms". In: 3rd International Conference on Interoperability for Enterprise Software and Applications (I-ESA 2007)
[132] Delgado M.; Agostinho C.; Malo P.; Jardim-Gonçalves R.; (2006). "A framework for STEP-based harmonization of conceptual models". 3rd International IEEE Conference on Intelligent Systems (IEEE-IS'06), Wetsminster, Jul 4-6 , 2006.
[133] INTEROP NoE (2006). Deliverable D.A3.3 - Semantic Annotation language and tool for Information and Business Processes. InterOP, 2006
[134] Delgado, M. (2008). "Harmonisation of STEP and MDA conceptual models using Model Morphisms". MSc thesis presented at Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, 2008
[135] Dauben, J. W.; Georg C. (1979). "His Mathematics and Philosophy of the Infinite". Boston: Harvard University Press, ISBN 978-0-691-02447-9.
[136] Bernstein, P. A. (2003). "Applying Model Management to Classical Meta Data Problems". First Biennial Conference on Innovative Data Systems Research, California, January 2003.
[137] Giunchiglia, F.; Yatskevich, M.; Shvaiko, P. (2007). "Semantic Matching: Algorithms and Implementation". J. Data Semantics 9: 1-38 (2007)
[138] Blaha, M.; Premerlani, W. (1996). "A catalog of object model transformations". Proceedings of WCRE, Volume 96, pp 87, 1996.
[139] Nicolis, G., Prigogine I. (1989). "Exploring Complexity: An Introduction". W. H. Freeman and Company, New York (1989)
[140] Wiederhold, G. (1992). "Mediators in the architecture of future information systems". IEEE Computer Magazine 25 (3): 38—49. doi:10.1109/2.121508 (March 1992).
[141] Foucault M. (1980). "Power/Knowledge". New York: Pantheon.
[142] Leonard D. (1999). "Wellsprings of Knowledge - Building and Sustaining the Sources of Innovation". Boston, MA: Harvard Business School Press.
[143] Matthews, K.; Harris, H. (2006). "Maintaining Knowledge Assets". In proceedings of the 1st World Congress on Engineering Asset Management (WCEAM) 11 – 14 July 2006 pp. 618-626.
[144] Metaxiotis K. (2005). "Exploring the world of knowledge management: agreements and disagreements in the academic/practitioner community". Journal of Knowledge Management, 9(2), 6-18.
[145] Witherell, P. (2009). "Semantic Methods for Intelligent Distributed Design Environments". In: Open Acess Dissertations of the ScholarWorks@UMass Amherst - University of Massachusetts - Amhers 2009.
[146] Berrar, D.; Sato, N.; and Schuster, A.; (2010). "Quo Vadis, Artificial Intelligence?". In: Advances in Artificial Intelligence - Hindawi Publishing Corporation, Volume 2010, Article ID 629869.
[147] Purves, D.; Augustine, G. J.; Fitzpatrick, D.; Lawrence, C. K.; Lamantia, A-S, McNamara, J. O.; and Williams, S. M. (2004). "Neuroscience", Sinauer Associates, Sunderland, Mass, USA, 3rd edition, 2004.
[148] Saha, D. and Mukherjee, A. (2003). "Pervasive computing: a paradigm for the 21st century," Computer, vol. 36, no. 3, pp. 25–31, 2003.
[149] Kephart, J. O. and Chess, D. M. (2003). "The vision of autonomic computing," Computer, vol. 36, no. 1, pp. 41–50, 2003.
[150] Roberta, F. (2009). "Reverse engineering the brain," Biomedical Computation Review, vol. 5, no. 2, pp. 10–17, 2009.
[151] Hylton, T. (2011). "DARPA: Defense Sciences Office". Retrieved from the web at September 2011: http://www.darpa.mil/Our_Work/DSO/Programs/Systems_of_Neuromorphic_Adaptive_Plastic_Scalable_Electronics_(SYNAPSE).aspx

[152] IBM (2011). “New ways of thinking”. Retrieved from the web at September 2011: http://www.ibm.com/smarterplanet/us/en/business_analytics/article/cognitive_computing.html
[153] Tulving, E. (1972). “Organisation of Memory” (eds Tulving, E. & Donaldson, W.) 381-403 (Academic, New York, 1972).
[154] Patterson, K.; Nestor, P. J.; and Rogers, T. T. (2007). “Where do you know what you know? The representation of semantic knowledge in the human brain”. In: Nature Reviews Neuroscience 8, 976-987 (December 2007).
[155] Koch C. (1999). “Biophysics of Computation. Information Processing in Single Neurons”. Oxford, UK: Oxford Univ. Press. 562 pp.
[156] Rolls ET, Treves A. (1998). “Neural Networks and Brain Function”. Oxford, UK: Oxford Univ. Press. 418 pp.
[157] Rolls ET. (2000). “Memory systems in the brain”. In: Annu Rev Psychol. 2000; 51:599-630
[158] Rolls ET. (1990). “A theory of emotion, and its application to understanding the neural basis of emotion”. Cogn. Emot. 4:161–90
[159] Rolls ET. (1999). “The Brain and Emotion”. Oxford, UK: Oxford Univ. Press. 367 pp.
[160] Rolls ET. (2000). “Précis of the brain and emotion”. Behav. Brain Sci. In press
[161] Wikipedia, (2010). “Hebbian Theory”. Retrieved from the web: http://en.wikipedia.org/wiki/Hebbian_theory, at August 2010.
[162] Stergiou, C.; and Siganos, D. (2008). “Neural Networks”. In: Imperial College of London, Computer Department; http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html 02-05-2010
[163] Simon S. Haykin (2008). In the book: “Neural networks and learning machines”, November 28, 2008 | ISBN-10: 0131471392 | ISBN-13: 978-0131471399 | Edition: 3.
[164] Orr, G.; Schraudolph, N.; and Cummins, F.; (1999). “CS-449: Neural Networks”. In: Willamette University Lecture Notes of the fall 99. Retrieved from the web at October 2011: http://www.willamette.edu/~gorr/classes/cs449/ann-overview.html
[165] Sherrod, P. (2011). DTREG - Software For Predictive Modeling and Forecasting. Retrieved from the web at December 2011: http://www.dtreg.com/
[166] Bombardier, V.; Mazaud, C.; Lhoste P.; and Vogrig, R. (2006). “Contribution of fuzzy reasoning method to knowledge integration in a defect recognition system”. In: Computer in Industry, In Press, 2006.
[167] Bulbul, A. (2010). “A Short Fuzzy Logic Tutorial”. Retrieved from the web at December 2011: http://www.cs.bilkent.edu.tr/~bulbul/depth/fuzzy.pdf
[168] Hüllermeier, E. (2011). “Fuzzy Logic in Machine Learning”. In a talk at Simpósio Brasileiro de Automação Inteligente SBAI-2011, São João del-Rei, Brasil, 2011. Retrieved from the web at October 2011: http://www.ppgel.net.br/sbai2011/arquivos/Talk.SBAI.11.pdf.
[169] Zhou, L. (2007). “Ontology learning: state of the art and open issues”. Published online: 24 March 2007 – Springer Science+Business Media, LLC 2007. Inf. Technol. Manage. (2007) 8:241-252 DOI 10.1007/s 10799-007-0019-5.
[170] Wikipedia (2011). “Operations Research”. Retrieved from the web: http://en.wikipedia.org/wiki/Operations_research, at November 2011.
[171] Bureau of Labor Statistics (2011). “Operations Research Analysts”. Retrieved from the web: http://www.bls.gov/oco/ocos044.htm, at November 2011.
[172] Craven, B. D. and Islam, SMN, (2006). “Operations Research Methods – Related Production, Distribution and Inventory Management Applications”. The Icfai University Press, Hyderabad, (2006).
[173] Oliveira, P.; and Gomes, P. (2009). “Probabilistic Reasoning in the Semantic Web using Markov Logic”. MSc Thesis - Knowledge and Intelligent Systems Laboratory - Cognitive and Media Systems Group - Centre for Informatics and Systems of the University of Coimbra
[174] Domingos, P.; Kok, S.; Lowd, D.; Poon, H.; Richardson, M.; and Singla, P. (2008). “Markov Logic,” Probabilistic Inductive Logic Programming, 2008, pp. 92-117.´
[175] Chen, J. (2009). “Markov Logic Network for Information. Extraction”. In: Sciences New York. Retrieved from the web site of the Department of Computer Sciences - University of Texas at Austin at May 2012: ftp://ftp.cs.utexas.edu/pub/techreports/honor_theses/cs-08-11-chen.pdf.
[176] Jalbuena, G.; Cory, J. (2010). Exact Sampling of Skew Young Diagrams. Retrieved from the web at December 2011: http://cnx.org/content/m34949/latest/
[177] Mollinari, M. (2007). “Aplicações das Cadeias de Markov na Genética, Seminários em Genética e Melhoramento de Plantas” – LGN 5799, 2007.
[178] Page, L. and Brin, S.; Motwani, R. and Winograd, T. (1999). “The PageRank Citation Ranking: Bringing Order to the Web”. Technical Report
[179] Albagli, S.; Ben-Eliyahu-Zohary, R.; and Shimony, S. E. (2009). Markov network based ontology matching. In: Proceedings of the 21st international Joint Conference On Artificial Intelligence - Pasadena, California, USA

[180] Drumond, L. (2009). “Aquisição automatizada de hierarquias de conceitos de ontologias utilizando aprendizagem estatística relacional”. MSc dissertation presented at: Universidade Federal do Maranhão, Programa de Pós-Graduação em Engenharia da Electricidade, 2009.
[181] Svátek, V. (2006). “Ontology Mapping enhanced using Bayesian Networks”, 2006.
[182] Chickering, D (2002). “Learning equivalence classes of bayesian-network structures”. Published in: The Journal of Machine Learning Research archive - Volume 2, 3/1/2002
[183] Zhou, H. (2002). “The code project - Inference in Belief Networks”. Retrieved from the web at December 2011: http://www.codeproject.com/KB/recipes/beliefnet.aspx?display=Print
[184] Harris, C. and Cahill, V. (2005). “Power management for stationary machines in a pervasive computing environment”. In Proceedings of the Hawaii International Conference on System Sciences, 2005.
[185] Arroyo-Figueroa, G. and Sucar, L. (1999). “A temporal bayesian network for diagnosis and prediction”. In Proceedings of the 15th Annual Conference on Uncertainty in Artificial Intelligence (UAI-99), pages 13–20, San Francisco, CA, 1999. Morgan Kaufmann Publishers.
[186] Devitt, A.; Danev, B. and Matusikova, K. (2006). Ontology-driven Automatic Construction of Bayesian Networks for Telecommunication Network Management, 2006.
[187] Colace, F.; and De Santo, M. (2010). “Ontology for E-Learning: A Bayesian Approach.” In: IEEE Transactions on Education, vol. 53, no. 2, May 2010.
[188] Ding, Z.; Peng, Y.; and Pan, R. (2004). “A Bayesian Approach to Uncertainty Modeling in OWL Ontology”. In Proceedings of 2004 International Conference on Advances in Intelligent Systems - Theory and Applications (AISTA2004) – November 2004. Luxembourg- Kirchberg, Luxembourg.
[189] Ding, Z.; and Peng, Y. (2004). “A Probabilistic Extension to Ontology Language OWL”. In Proceedings of the 37th Hawaii International Conference on System Sciences (HICSS-37), January 2004. Big Island, Hawaii.
[190] Peng, Y.; and Ding, Z. (2005). “Modifying Bayesian Networks by Probability Constraints”. In Proceedings of 21st Conference on Uncertainty in Artificial Intelligence, Edinburgh, Scotland, July 26-29, 2005.
[191] Ding, Z. (2005). “BayesOWL: A Probabilistic Framework for Semantic Web”. In a PhD Thesis presented at Computer Science and Electrical Engineering Department of the University of Maryland, Baltimore County – 2005.
[192] The PR-OWL/MEBN/UnBBayes team (2010). “PR-OWL: A Bayesian extension to the OWL Ontology Language”. Retrieved from the web at November 2011: http://www.pr-owl.org/.
[193] Costa, P. C. G.; Laskey, K. B.; Laskey, K. J. (2003). “PR-OWL: A Bayesian Ontology Language for the Semantic Web”. In Center for Technology-Enhanced Learning, University of Missouri-Rolla (2003).
[194] Druzdzel, M. J.; Van der Gaag, L. C. (2000). “Building bayesian networks: Where do the numbers come from?”. In: IEEE Transactions on Knowledge and Data Engineering, 12:481–486, 2000.
[195] Wikipedia, (2011). “Machine Learning”. Retrieved from the web: http://en.wikipedia.org/wiki/Machine_learning, at August 2011.
[196] Schapire, R. (2008). COS 511: Theoretical Machine Learning. In Princeton courses. Retrieved from the web at June 2011: http://www.cs.princeton.edu/courses/archive/spr08/cos511/scribe_notes/0204.pdf
[197] Witten, I.; Frank, E. (2005). “Data mining: practical machine learning tools and techniques”. Second Edition (Morgan Kaufmann Series in Data Management Systems). ISBN-13: 978-0-12-088407-0 | 978-0120884070 | 9780120884070; ISBN-10: 0-12-088407-0 | 0120884070.
[198] Sewell, M. (2007). “Machine Learning”. Retrieved from the web at October 2011: http://machine-Learning.martinsewell.com/machine-Learning.pdf
[199] Jones & Berlett Learning (2011). “Classification Learning” Chapter. Retrieved from the web at November 2011: http://samples.jbpub.com/9780763785871/85871_CH03_FINAL.pdf
[200] Keller, F. (2001). “Connectionist and Statistical Language Processing”. (Lecture Slides). Computerlinguistik, Universitat des Saarlandes, Saarbrucken, 2001.
[201] Dijkstra, E. W. (1959). "A note on two problems in connexion with graphs". Numerische Mathematik 1: 269–271, 1959
[202] Young, N. E. (2004). “ShortestPathsByDijkstras”. Retrieved from the web at November 2011: http://www.cs.ucr.edu/~neal/2004/cs141/?ShortestPathsByDijkstras.
[203] Agrawal, R.; Imielinski, T.; Swami, A. N. (1993). “Mining Association Rules between Sets of Items in Large Databases”. In Proc. of the ACM SIGMOD Intl. Conf. on Management of Data (SIGMOD’93), pages 207-216, Washington D. C.
[204] Hipp, J.; Güntzer, U.; Nakhaeizadeh, G. (2000). “Algorithms for Association Rule Mining—A General Survey and Comparison,” SIGKDD Explorations, vol. 2, no. 1, pp. 58-64, July 2000

[205] Wikipedia (2011). “Association Rule Learning”. Retrieved from the web at November 2011: http://en.wikipedia.org/wiki/Association_rule_learning

[206] Data privacy Lab (2011). Learning Robust Association Rules from Data. Retrieved from the web at November 2011: http://dataprivacylab.org/dataprivacy/projects/gentree/index.html

[207] Auslender, L. E. (2004). “On analytical tools for market basket (associations) analysis”. Presented at NYC Informs, New York City, May 2004. Retrieved from the web at November 2011: http://nymetro.chapter.informs.org/prac_cor_pubs/Ausleder-On-market-basket-analysis-May-04.pdf.

[208] Agrawal, R.; Srikant, R. (1994). “Fast algorithms for mining association rules”. In: Proceedings of the 20th VLDB conference, pp 487–499.

[209] Malik, J. S.; Goyal, P.; Sharma, A. K. (2011). “Data Mining Problem Solving Algorithms & Their Comparative Study”. Presented at Pioneer Institute of Professional Studies, Indore - 12th National Conference On ETCC – 2011: Emerging Trends in Computing & Communication (May 6, 2011)

[210] Wu, X.; Wu, V. Kumar, V.; J.R. Quinlan, J. R. ;, J. Ghosh, J.; Q. Yang, Q.; H. Motoda, H.; G.J. McLachlan, G. J.; A.F.M. Ng, A. F. M.; B. Liu, B.; P.S. Yu P. S. , Z.-H. Zhou Z.-H.;, M. Steinbach, M.; D.J. Hand, D. J.; and D. Steinberg, D. (2008). "Top 10 Algorithms in Data Mining," *Knowledge and Information Systems,* vol. 14, no. 1, pp. 1-37, 2008.

[211] Spehr, J.; Winkelbach, S. (2011). “The k-means algorithm”. Accessed at November 2011: http://www.rob.cs.tu-bs.de/content/04-teaching/06-interactive/Kmeans/Kmeans.html

[212] Minka, T. (2001). In lecturers about data mining (2001) - Day 31 - Logistic regression. Retrieved from the web at November 2011: http://alumni.media.mit.edu/~tpminka/courses/36-350.2001/lectures/day31/

[213] Met Office (2011). Using numerical models to create weather forecasts. Retrieved from the web at November 2011: http://www.metoffice.gov.uk/learning/science/days-ahead/models

[214] Met Office (2011). Conservative semi-Lagrangian transport schemes. Retrieved from the web at November 2011: http://www.metoffice.gov.uk/research/areas/dynamics/conservative-schemes.

[215] Gartner (2011). “Gartner IT Glossary”. Retrieved from the web: http://www.gartner.com/technology/it-glossary/, at July 2011.

[216] Slidefinder (2011). Retrieved from the web at March 2011: http://www.slidefinder.net/D/Data_Mining_Timo_Knuutila_Department/994490

[217] Wikipedia 2011. Retrieved from the web at November 2011: http://en.wikipedia.org/wiki/Markov_random_field

[218] Srinidhi B. (1998) “Needed: agile accounting to match agile organisations”. Agility & Global Competition, 2(1), 41-55.

[219] Ashkenas RN. (1995). “The boundaryless organization: breaking the chains of organizational structure”. San Francisco: Jossey-Bass Publishers.

[220] Goldman SNRPK. (1995). “Agile Competitors and Virtual Organizations: Strategies for Enriching the Customer”. New York: Van Nostrand Reinhold.

[221] Goldman SL, Nagel RN & Priess K. (1995). “Agile Competitors and Virtual Organizations: Strategies for Enriching the Customer”. New York: Van Nostrand Reinhold.

[222] Bessant, J.; Francis, D.; Meredith, S.; Kaplinsky, R.; Brown, S. (2001). “Developing manufacturing agility in SMEs”. International Journal of Technology Management, (22), 28-54.

[223] Gadomski, A.M. (1986). “Application of System-Process-Goal Approach for description of TRIGA RC1 System”. In: 9th. European TRIGA Nuclear Reactor Users Conference ", October 1986.

[224] Smith, M. K. (2003). “Learning theory”. In the encyclopedia of informal education, www.infed.org/biblio/b-learn.htm, Last update: December 04, 2011.

[225] Ramsden, P. (1992). “Learning to Teach in Higher Education”, Book published in 1992 by Routledge - 290 + xiv pages.

[226] Stockley, D. (2010). “E-Learning definition and explanation”; Available at: http://derekstockley.com.au/elearning-definition.html; Last access on February 2010.

[227] Nielson, B. (2009). “Relevance of Learning versus Relevance of Training and Development”. Available at: http://ezinearticles.com/?Relevance-of-Learning-Versus-Relevance-of-Training-and-Development&id=2255045; Last access on March 2010.

[228] Sarraipa J.; Figueiredo, D.; Maló, P.; and Jardim-Gonçales, R. (2010). “An Inter-Organisational approach to Industrial e-Training”; The Summer 4th International Conference on Knowledge Generation, Communication and Management: KGCM 2010;

[229] Nicholson, P. (2005); “E-training or E-learning? Towards a synthesis for the knowledge-era workplace”; International Federation for Information Processing; 2005.

[230] Sarraipa, J; Malo, P.; Patel, H.; Pettitt, M.; Hardiman, S.; Banassino, M.; Syllignakis, A; Carter, M.; (2009). “D102 -Training Strategy and Plan”. A CoSpaces project deliverable. 30th April 2009

[231] Smith, R. S. (2004). “Guidelines for authors of learning objects”. 2004 NMC: The New Media Consortium. Retrieved from the web at October 2011: http://archive2.nmc.org/guidelines/NMC%20LO%20Guidelines.pdf
[232] Dick, W.; Carey, L.; Carey, J. O. (2004). “The Systematic Design of Instruction”. ISBN-10: 0205412742 | ISBN-13: 978-0205412747 | Publication Date: July 29, 2004 | Edition: 6
[233] Maló P.; Sarraipa J.; Jardim-Gonçalves R.; Steiger-Garção A. (2008). “The CoSpaces Training System”. In: ICE2008 Conference Proceedings - 14th International Conference on Concurrent Enterprising, 23-25 Jun 2008, Lisbon, Portugal.
[234] BusinessDictionary.com (2009). Training Programme definition; Available at: http://www.businessdictionary.com/definition/training-program.html; Last accessed on January 2009
[235] Sarraipa, J; Malo, P.; Costas; Vassilis; Carter, M.; (2007). “D65 -Training Strategy and Plan”. A CoSpaces project deliverable. 18 December 2007
[236] IEEE (2002). “Draft Standard for Learning object metadata” - IEEE 1484.12.1-2002. Institute of Electrical and Electronics Engineers, Inc.; 2002.
[237] Kumar A. (2009). “Specifications and Standards in E-Learning”. Available at CDAC Mumbai Website: http://www.cdacmumbai.in/design/corporate_site/override/pdf-doc/Specifications_and_Standards_in_E-Learning.pdf; Retrieved at February 2010.
[238] Romessis, C.; Johansson, C.; Yves, B; Sarraipa, J. (2011). ‘BDA On-line training Portal providing access to training material and project results’, 2011, CRESCENDO Deliverable D1.5.4., R1.0.
[239] UVA-WEBLEARN - Your Online Learning Resource (2011). Blackboard Software Review. Available at: http://uva-weblearn.net/blackboard_review.asp; Retrieved at February 2011.
[240] Moodle 2011. Moodel about. Retrieved from the web at November 2011: http://moodle.org/about/
[241] Santos, L. (2009). “Adaptability Support in a Learning Management System”. Dissertation submitted to obtain a MSc degree in Informatics by Faculdade de Ciências da Universidade de Lisboa in December 2008.
[242] Braxton, S.; Bronico, K.; and Looms, T. (1995). “Instructional design methodologies and techniques”. (Dead web page)
[243] Edutech Wiki (2011). “Instructional design model”. Retrieved from the web page at December 2011: http://edutechwiki.unige.ch/en/Instructional_design_model.
[244] Cooper, S. (2009). “Theories of Learning in Educational Psychology”. In: Learning Theories Site Map. Retrieved from the web at December 2011: http://www.lifecircles-inc.com/Learningtheories/learningmap.html
[245] Training and Development.Naukrihub.com (2007). “Methods of Training”. Retrieved from the web at December 2011: http://traininganddevelopment.naukrihub.com/methods-of-training/
[246] McGriff, Steven (2001). “Instructional Systems Design Models”. Pennsylvania State University. (Dead web page).
[247] Merriënboer, J. J. G. (1997). “Training Complex Cognitive Skills: A Four-Component Instructional Design Model for Technical Training”. Englewood Cliffs, NJ: Educational Technology Publications.
[248] Clark, D. (2010).” Why Instructional System Design?”. Retrieved from the web at December 2011: http://www.nwlink.com/~donclark/hrd/sat1.html
[249] Molenda, M (2003). “The ADDIE Model”. In the draft submitted for publication in A. Kovalchick & K. Dawson, Ed's, Educational Technology: An Encyclopedia. Copyright by ABC-Clio, Santa Barbara, CA, 2003. Retrieved from the web at December 2011: http://www.indiana.edu/~molpage/The%20ADDIE%20Model_Encyclo.pdf
[250] Ira C. Eaker Center’s Public Website (2005). “Instruction to Instructional Systems Development”. In: Teaching and learning Theories materials. Retrieved from the web at November 2011: http://www.au.af.mil/au/awc/awcgate/doe/isd/paper.htm
[251] Learning-Theories.com (2010). “ADDIE Model”. In: Knowledge Bae and Webliography. Available at: http://www.learning-theories.com/addie-model.html; Retrieved at March 2010.
[252] Brinklow, T. (2004). “Domains, Ontologies, Models and the Knowledge Creation Cycle”. Published at Brighton Business School - Occasional / Working paper series - BBSW04-3; 2004.
[253] Ackoff, R. L. (1989). "From Data to Wisdom", Journal of Applies Systems Analysis, Volume 16, 1989 p 3-9.
[254] Keller, T.; and Tergan, S.O. (2005). “Visualizing Knowledge and Information: An Introduction”. In Knowledge and Information Visualization: Searching for Synergies. Springer 2005: p. 1-23.
[255] Wikipedia (2010). “Explicit Knowledge”. Retrieved from the web at August 2010: http://en.wikipedia.org/wiki/Explicit_knowledge.

[256] United States Army (2008). In annex A: Glossary of Terms of the “Army Knowledge Management Principles“ publication. Retrieved from the web at August 2010: http://www.army.mil/ciog6/docs/AKMPrinciples.pdf
[257] Willcocks, L. P. and Finnegan, D. (2007). “Implementing CRM - from technology to knowledge”. John Wiley series in information systems. Wiley, Chichester, UK. ISBN 9780470065266.
[258] Diamantini, C. (2009). “Is ‘Understanding’ a Misunderstanding?” In proceedings of the 15th International Conference on Concurrent Enterprising, Leiden – The Nerthelands, 22-24 June 2009.
[259] Halevy, A. Y. (2003). “Data Integration: A Status Report”. In: Datenbanksysteme für Business, Technologie undWeb. Leipzig Germany, February 2003.
[260] Lenzerini, M. (2002). “Data Integration: a Theoretical Perspective”. In: ACM Symposium on Principles of Database Systems,, 2002, p. 233-246.
[261] Missikoff, M., D’Antonio, F., Lenzerini, M., Jeusfeld, M., Johannesson, P. (2006). “Enterprise Modelling and Ontologies for Interoperability”. INTEROP project, FP6-IST-2002-508011, Deliverable D8.2, 2006.WWW page http://interop-vlab.eu/ei_public_deliverables/interop-noe-deliverables/do-ontology-for-interoperability.
[262] NIST (1993). “Integration Definition for Function Modeling (IDEF0): Draft”. In: Federal Information Processing Standards Publications No 183. Gaithersburg, Md, USA: National Institute for Standards and Technology.
[263] The European Future Internet Initiative (2010). “The European Future Internet Initiative (EFII)”. In: White paper on the Future Internet PPP Definition, January 2010.
[264] Rocket Project WP2 Partners (2002). “Deliverable D 2.1 State-of-the-Art of Knowledge Management”, from project ROCKET - Roadmap to Communicating Knowledge Essential for the indusTrial environment - IST-2001-38245.
[265] Miah, S. J.; Gammack, J.; Kerr, D. (2007). “Ontology development for context-sensitive decision support”. In proceedings: Third International Conference on Semantics, Knowledge and Grid – SKG2007 Xi'an, China, 29-31 October 2007 - Publisher IEEE.
[266] Staab, S.; Schnurr, H.-P.; Studer, R.; and Sure, Y.; (2001). “Knowledge processes and ontologies”. IEEE Intelligent Systems, Special Issue on Knowledge Management, 16(1), January/Febrary 2001.
[267] Drew O. M. & Duncan H. Br. (2001). “Section one – Manual for Media Trainers: A learned-Centred Approach”. Published by Asia-Pacific Institute for Broadcasting Development (AIBD), Malaysia and UNESCO.
[268] Levine, D.; Gilbert, A. (1999). “Managerial Practices Underlying One Piece of the Learning Organization”. In Knowledge Transfer website of Institute of Industrial Relations, University of California, Berkeley. Retrieved from the web at January 2011: http://www.irle.berkeley.edu/cohre/knowledge.html
[269] Smith, M. K. (1996, 2000). “Curriculum theory and practice”. In: The encyclopedia of informal education, www.infed.org/biblio/b-curric.htm. Last updated: 11 April 2008.
[270] Nonaka, I. and Konno, N. (1998). ‘‘The concept of ‘ba’: building a foundation for knowledge creation’’, California Management Review, Vol. 40 No. 3, pp. 40-54.
[271] McLean, L. (2012). “A Review and Critique of Nonaka and Takeuchi’s Theory of Organizational Knowledge Creation”. Retrieved from the web at April 2012: http://mcleanglobal.com/public/MGC/publications/Nonaka%20and%20Takeuchi.pdf.
[272] Geytere, T. (2012). SECI model (Nonaka Takeuchi). Retrieved from the web at April 2012: http://www.12manage.com/methods_nonaka_seci.html.
[273] Kaiser, A; Fordinal, B. (2010). "Creating a ba for generating self-transcending knowledge", Journal of Knowledge Management, Vol. 14 Iss: 6, pp.928 – 942.
[274] Assudani, R. H.(2005). "Catching the chameleon: understanding the elusive term “knowledge”", Journal of Knowledge Management, Vol. 9 Iss: 2, pp.31 – 44
[275] Scharmer, C. O. (2001). "Self-transcending knowledge: sensing and organizing around emerging opportunities", Journal of Knowledge Management, Vol. 5 Iss: 2, pp.137 – 151
[276] Nonaka, I.; Konno, N.; Toyama, R. (2001). “Emergence of ‘Ba’”. In: Knowledge Emergence: Social, Technical, and Evolutionary Dimensions of Knowledge Creation, pp 13 - 29, Oxford University Press US, (2001).
[277] IEEE Computer Society, New York –USA, (1995). IEEE Guide for Software Quality Assurance Planning. Std. 730.1-1995.
[278] Mika, P.; Akkermans, H. (2003). “D1.2 Analysis of the State-of-the-Art in Ontology-based Knowledge Management”. In: SWAP EU IST-2001-34103 Project Deliverable D1.2 (WP1).
[279] Donini, F. M.; Lenzerini, M.; Nardi, D.; Schaerf, A. (1991). “A hybrid system with datalog and concept languages: Trends in Artificial Intelligence”. Volume LNAI 549, Springer Verlag.

[280] Maedche, A.; Motik, B; Stojanovic, L. Studer, R.; Volz, R. (2003). “Ontologies for Enterprise Knowledge Management”. In: Journal IEEE Intelligent Systems, IEEE.
[281] Kalfoglou, Y.; Schorlemmer, M. (2003). “Ontology mapping: the state of the art”. Knowl. Eng. Rev., 18(1):1--31, 2003.
[282] Tsarkov, D.; Horrocks, I. (2003). “DL reasoner vs. first-order prover”. In Proc. of the 2003 Description Logic Workshop (DL 2003), volume 81 of CEUR (http://ceur-ws.org/), pages 152-159.
[283] Matzakou, I.; Askounis, D.; Sarraipa, J., and Jardim-Goncalves, R. (2012). “Enhanced Knowledge Management of an Interoperability Service Utility”. In the proceedings of the 6th IEEE International Conference on Intelligent Systems IS’12, Sofia, Bulgaria, September 6-8, 2012
[284] Noy, N.; McGuinness, D. (2002). “Ontology Development 101: A Guide to Creating Your First Ontology”. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford MIT-R SMI-2001-0880.
[285] Jardim-Goncalves, R.; Olavo, R.; Sarraipa, J.; Steiger-Garcao, A. (2004). “Ontology-Based Framework for Enhanced Interoperability in Networked Industrial Environments”. In: 11th IFAC Symposium on Information Control Problems in Manufacturing - INCOM 2004.
[286] Jardim-Goncalves, R.; Silva, J. P.M.A.; Steiger-Garcao, A.; Monteiro, A. (2007). “Framework for Enhanced Interoperability Through Ontological Harmonization of Enterprise Product Models in book Ontologies: A Handbook of Principles, Concepts and Applications”. Integrated Series in Information Systems , Vol. 14-2007, XXIV, 915 p., 245 illus. ISBN: 978-0-387-37019-4.
[287] Sarraipa, J.; Zouggar, N.; Chen, D; Jardim-Goncalves, R. (2007). “Annotation for Enterprise Information Management Traceability”. In Proceedings of IDETC/CIE ASME 2007
[288] Agostinho, C. Correia, F. and Jardim-Goncalves, R. (2010). “Interoperability of Complex Business Networks by Language Independent Information Models”. In: 17th ISPE International Conference on Concurrent Engineering (CE 2010). Sep 6-10, Cracow, Poland
[289] Agostinho, C., Černý, J., & Jardim-goncalves, R. (2012). “MDA-Based Interoperability Establishment Using Language Independent Information Models”. In M. van Sinderen, P. Johnson, X. Xu, & G. Doumeingts (Eds.), 4th International IFIP Working Conference (IWEI 2012) (pp. 146–160). Harbin, China: Springer. Retrieved at September 2012 from http://www.springerlink.com/content/p7582p4485132643/
[290] ISO/IEC (2004). ISO/IEC: Information Technology - Metadata registries (MDR) - Part 1: Framework (ISO/IEC 11179-1:2004), (2004).
[291] Jouault F, Kurtev I, (2007). “On the interoperability of model-to-model transformation languages”. Science of Computer Programming, no. 68, pp. 114–137.
[292] Grangel R., Bigand M., Bourey J.P. (2008). “A UML Profile as Support for Transformation of Business Process Models at Enterprise Level”. In: MDISIS 2008
[293] Wimmer M, Seidl M. (2009). “Model Transformation with ATL”. In: 1st International Workshop, France, July 8-9, MtATL
[294] Jardim-Goncalves, R.; Agostinho, C.; Sarraipa, J.; Grilo, A.; and Mendonça, J. (2012). “Reference Framework for Enhanced Interoperable Collaborative Networks in Modern Furniture Industrial Organizations: The funStep Case Study”. Submitted in 2012 and published in:· International Journal of Computer Integrated Manufacturing - Collaborative Networks as Modern Industrial Organisations: Real Case Studies archive - Volume 26 Issue 1-2, January-February 2013 - Pages 166-182.
[295] Brunnermeier, S. and Martin S.A. (1999). Interoperability Cost Analysis of the U.S. Automotive Supply Chain: Final Report. DIANE Publishing.
[296] Ray, S.R. and Jones, A.T. (2006). “Manufacturing interoperability”. Journal of Intelligent Manufacturing, vol.17 (6), pp. 681-688.
[297] Jardim-Goncalves, R., Figay, N., Steiger, A.(2006). “Enabling interoperability of STEP Application Protocols at meta-data and knowledge level”. International Journal of Technology Management, Vol. 36, N 4 / 2006 , pp. 402 – 421.
[298] Molina, A., Rodriguez, C. A., Ahuett, H., Cortés, J. A., Ramírez, M., Jiménez, G. and Martinez, S.(2005). “Next generation manufacturing systems: key research issues in developing and integrating reconfigurable and intelligent machines”. In: International Journal of Computer Integrated Manufacturing, 18:7,525 — 536.
[299] IMS Glossary (2009). “Intelligent” Definition. Retrieved from the web (http://www.ims.org/content/glossary) in June 2009
[300] ISO. (2006). ISO 10303-236:2006. Industrial automation systems and integration—product data representation and exchange—Part 236: Application protocol: Furniture catalog and interior design.
[301] Breiter, A. and Light, D. (2004). “Decision Support Systems in Schools – from Data Collection to Decision Making”. In: AMCIS 2004 -Tenth Americas Conference on Information Systems, August 2004, New York, USA.
[302] Pustejovsky, J., Ed. (1993). “Semantics and the lexicon”. Dordrecht, Kluwer Academic Press.

[303] Sendall, S., Strohmeier, A. (2001). Requirements Analysis with Use Cases, v1.0. In: SWEED 2001. Retrieved from the web at October 2009: http://www.uml.org.cn/requirementproject/pdf/re-a2-theory.pdf
[304] Willerval, B., Boisseau, Y., Serres-Cousiné, H. (Eds) (1989). Le petit Larousse illustré. Librairie Larousse.
[305] Boudjlida N., Panetto H. (2007). "Enterprise Semantic Modelling for Interoperability". Proceedings of 12th IEEE Conference on Emerging Technologies and Factory Automation (ETFA'2007). 847-854, September 25-28, Patras, Greece, ISBN: 978-1-4244-0826-9
[306] Boudjlida, N., Panetto, H., Baïna, S., Diamantini, C., Krogstie, J., Lin, Y., Sarraipa, J., Zouggar, N., Hahn, A., Delgado, M., Abian, M.-A., Nunez, M.J., (2007). INTEROP deliverable "DTG4.2: Experimental Semantic Enrichment of Enterprise Models for Interoperability and its Practical Impact". EC project IST 508-011.
[307] Gatlin, K.A. (2005). "Enhancing Cross-Language Retrieval of Comparable Corpora Through Thesaurus-Based Translation and Citation Indexing". In UNC SILS Master's Papers. School of Information and Library Science.
[308] Ballesteros, L. and Croft, B. (1996). "Dictionary methods for cross-lingual information retrieval". In: DEXA '96 - Database and Expert Systems Applications Conference, September 9-13, Zurich, Switzerland.
[309] Soergel, D. (1997). "Multilingual thesauri in cross-language text and speech retrieval". In: AAAI Symposium on Cross-Language Text and Speech Retrieval, March 24-26, Stanford, California, USA.SMART-fm. (2004). A standards-compliant framework to support complete integrated product life-cycle information management and electronic commerce for the furniture manufacturing (fm) industry, in the advent of the smart enterprises. EC project No. IST-2001-52224.
[310] Jardim-Gonçalves, R., Agostinho, C., Malo, P., and Steiger-Garcao, A. (2007). "Harmonising technologies in conceptual models representation". Int. J. Product Lifecycle Management, vol.2 (2), pp. 187 – 205.
[311] INNOVAFUN. (2007). "Deliverable 1.1: Opportunities for Innovation" [online]. [Accessed 12 November 2008] Available from World Wide Web: http://standards.eu-innova.org/Files/Report/D1.1.Opportunities_for_innovation_using_funStep.pdf. EC INNOVA Project No.: 031139.
[312] Brewster, C. (2010). Ontology Learning. In Ontogenesis - An Ontology Tutorial – retrieved from the web at August 2010: http://ontogenesis.knowledgeblog.org/331.
[313] Gómez-Pérez, A.; Staab, S.; and Studer, R. (2004) "Ontology Evaluation". In: Handbook on Ontologies, Springer, 2004, 251-274
[314] Oltramari, A.; Gangemi, A.; Guarino, N. & Masolo, C. (2009). "Restructuring WordNet's Top-Level: The OntoClean approach". Proceedings of the Workshop OntoLex'2, Ontologies and Lexical Knowledge Bases, 2002
[315] Ratliff, N.; Silver, D. (2009). "Learning to Search: Functional Gradient Techniques for Imitation Learning". In Journal Autonomous Robots; Publisher: Springer Netherlands; 2009-07-01; Issn: 0929-5593; Issue 1; Pages 25-53
[316] Wikipedia 2012. Definition of "Google Search". Retrieved from the web at September 2012: http://en.wikipedia.org/wiki/Google_Search.
[317] Jardim-Goncalves, R., Agostinho, C., & Steiger-Garcao, A. (2012). "A reference model for sustainable interoperability in networked enterprises: towards the foundation of EI science base". International Journal of Computer Integrated Manufacturing, (February), 1–19. doi:10.1080/0951192X.2011.653831
[318] Wooldridge, M., (2009). "An Introduction to MultiAgent Systems". John Wiley and Sons.
[319] Wooldridge, N., Jennings, M., (1995). "Intelligent Agents: Theory and Practice". The Knowledge Engineering Review, vol. 10:2, pp. 115-152.
[320] Bellifemine, D., Caire, F., (2007). "G. Greenwood: Developing Multi-Agent Systems with JADE". Wiley.
[321] Ferreira, J., Agostinho, C., Sarraipa, J., Jardim-Goncalves, R., (2011). "Monitoring Morphisms to Support Sustainable Interoperability of Enterprise Systems". In 6th Int. Workshop EI2N'2011 of OTM Conference, Greece.
[322] Pallot, M. (2010). (CoSpaces Newsletter Editor). 1st CoSpaces Newlstter. CoSpaces Consortium. Available at: http://www.cospaces.org/downloads/CoSpaces%20Newsletter%201.pdf Retrieved : 22/01/2010
[323] Wiig, Karl M. (1999). "The Intelligent Enterprise and Knowledge Management". In: UNESCO's Encyclopedia of Life Support Systems. Knowledge Research Institute, Inc.
[324] Schmidt, A. (2008). "Enabling Learning on Demand in Semantic Work Environments: The Learning in Process Approach". IGI Publising.

[325] Camarinha-Matos, L.M.; Afsarmanesh, H. & Ollus, M. (2005). “ECOLEAD: A Holistic Approach to Creation and Management of Dynamic Virtual Organizations”. Collaborative Networks and Their Breeding Environments, pp. 3-16, Springer Science Publisher.
[326] Hamburg, I ; Engert, S. ; Thig, H. ten . (2007). “Improving Participation and Collaboration in Business-Oriented e-Learning for SME’s”. In: IFAC/ International Federation of Automatic Control, Seoul, 2007.
[327] ARIEL - Analysing and Reporting on the Implementation of Electronic Learning in Europe Consortium (2010). Analyzing and Reporting on the Implementation of Electronic Learning; Available at: http://www.ariel-eu.net; Last accessed on: February 2010.
[328] Hamburg I., Engert S., Anke P. and Marin M. (2008). “Improving e-Learning 2.0 based training strategies of SME’s through communities of practice”; In: SIMPEL final conference 2008.
[329] Ullrich C. (2004). “Description of an Instructional Ontology and its Application in Web Services for Education”. In: Proc. Workshop Applications of Semantic of Web Technologies for E-Learning; 2004.
[330] Gasevic D., Jovanovic J., Devedzic V. and Boskovic M.(2005). “Ontologies for Reusing Learning Object Content”; In: Fifth IEEE International Conference on Advanced Learning Technologies, 2005.
[331] Merceron A., Oliveira C., Scholl M. and Ullrich C. (2004). “Mining for Content Re-Use and Exchange – Solutions and Problems”; In: Poster Proceedings of the 3rd International Semantic Web Conference, 2004.
[332] Allert H., Markannen H. and Richter C. (2006). “Rethinking the Use of Ontologies in Learning”. In: Proc. 2nd International Workshop on Learner Oriented Knowledge Management and KM-Oriented Learning; 2006.
[333] Hurteau, M.; Houle, S., & Mongiat, S. (2009). "How Legitimate and Justified are Judgments in Program Evaluation?". Evaluation 15 (3): 307–319. doi:10.1177/1356389009105883.
[334] Oxford Online Dictionary (2012). “Consistency” Definition, available from http://oxforddictionaries.com/definition/consistency, accessed on February 2012.
[335] Haase, P., Harmelen, F., Huang, Z., Stuckenschmidt, H. and Sure, Y. (2005). “A Framework for Handling Inconsistency in Changing Ontologies”. Proceedings of the 4th International Semantic Web Conference (ISWC), November 6-10, Galway, Ireland, pp. 353-367, (2005).
[336] Baclawski, K., Kokar, M., Waldinger, R. and Kogut, P. (2002). “Consistency Checking of Semantic Web Ontologies”. In First International Semantic Web Conference (ISWC 2002), Sardinia.
[337] University of Oxford, Information Systems Group (2012). Hermit OWL Reasoner Overview, available from http://www.hermit-reasoner.com. Accessed on February 2012.
[338] Li, D., Huang, L. and Li, M. (2004). “Dynamic Semantic Consistency Checking of Multiple Collaborative Ontologies in Knowledge Management System”. Proceedings of the 5th international conference on Parallel and Distributed Computing: applications and Technologies (PDCAT), December 8-10, Singapore, pp. 76-80, 2004.
[339] Silva, J.; Monteiro, A.; Sarraipa, J.; Gonçalves, R. (2009). “Towards Ontology Harmonization of Mechanical Manufacture Constraints Through PLC”. In Proceedings 20th International Congress of Mechanical Engineering (COBEM 2009), pp 15 - 20, (2009).
[340] Jardim-Goncalves, R.; Bonfatti, F. (2008). “The Road for Seem: A Reference Framework Towards a Single European Electronic Market”. Taylor & Francis Group, London, UK.. Published by Press/Balkema. ISBN 978-0-415-41995-6.
[341] White W, O’Connor A, Rowe B, (2004). “Economic Impact of Inadequate Infrastructure for Supply Chain Integration”. NIST Planning Report 04-2, Gaithersburg, MD: National Institute of Standards and Technology.
[342] CoSpaces Consortium (2008). “Course EFB – Enabling new Forms of Business”. Retrieved from the web at May 2012: http://www.cospaces.org/training/training_courses.htm
[343] Rebelo de Sousa, António, (2010). “Gestão Financeira”. Lisboa: Instituto Superior de Ciências Sociais e Políticas, ISBN: 978-989-646-043-3
[344] CRESCENDO Consortium (2012). CRESCENDO introduction message. Retrieved from the web at May 2012: http://www.crescendo-fp7.eu/
[345] Graignic, P.; Quenardel, A.; Vosgien, T.; Tuloup, V.; Berquet, J.; Broeze, H.; Dubourg, G.; Gevers, A.; Armstrong, C; Nolan, D.; Richard, C.; Sarraipa, J. (2012). In the presentation: Product Integration process using innovative integration-ability in a collaborative environment. Presented at CRESCENDO Project Forum, 19-21 June 2012, Toulouse, France.
[346] Nekvasil M. (2010). “Evaluation of Semantic Applications for Enterprises”, dissertation thesis, VŠE-FIS, Prague, 2010
[347] Gallaher, M. P.; O’Connor, A. C.; Dettbarn, J. L., Jr.; Gilday, L. T. (2004). “Cost Analysis of Inadequate Interoperability in the U.S. Capital Facilities Industry”. August 2004. National Institute of Standards & Technology. GCR 04-867

[348] AccountingCoach (2012). Activity Based Costing. Retrieved from the web at May 2012: http://www.accountingcoach.com/online-accounting-course/35Xpg01.html
[349] Knoow.net (2012). Ciências Económicas e Empresariais Retrieved from the web at May 2012: http://www.knoow.net/cienceconempr/gestao/seissigma.htm)
[350] Mittal Consultants and Enterprises (2012).Six Sigma Calculation. Retrieved from the web at May 2012: http://www.miconleansixsigma.com/six-sigma-calculation.html
[351] Oh, Lih-Bin; Ng, Benjamin Loong-Tatt,; Teo, Hock-Hai; (2007). “IT Portfolio Management: A Framework For Making Strategic IT Investment Decisions”. In: ECISUniversity of St. Gallen (2007) , p. 1265-1275.
[352] Grilo, A., Zutshi, A., Jardim-Gonçalves, R., Steiger-Garção, A., (2012). “Construction collaborative networks: the case study of a building information modelling-based office building project”. International Journal of Computer Integrated Manufacturing. iFirst.
[353] Guédria,W.; Naudet, Y.; Chen, D. (2011). “Maturity Model as Decision Support for Enterprise Interoperability”. OTM Workshops 2011: 604-608.
[354] ISO. (2011). ISO 11354-1:2011. Advanced automation technologies and their applications -- Requirements for establishing manufacturing enterprise process interoperability -- Part 1: Framework for enterprise interoperability
[355] Alonso, J., De Soria, I., Orue-Echevarria, L., Vergara, M. (2010). “Enterprise Collaboration Maturity Model (ECMM): Preliminary Definition and Future Challenges”. In: Enterprise Interoperability IV, Part VII, pp. 429–438. Springer, Heidelberg
[356] Turnitsa, C.; Tolk, A. (2008). “Knowledge Representation and the Dimensions of a Multi-Model Relationship”. In Proceedings of the 40th Winter Simulation Conference, 2008: 1148-1156.
[357] Tolk, A., S. Y. Diallo, C. Turnitsa, L. S. Winters. (2006). “Composable M&S Web Services for Net-centric Applications”. Journal Defense Modeling and Simulation 3 (1): 27-44.
[358] Gannon, T.; Madnick, S.; Moulton, A.; Siegel, M.; Sabbouh, M.; Zhu, H.; (2009). “Framework for the Analysis of the Adaptability, Extensibility, and Scalability of Semantic Information Integration and the Context Mediation Approach”. Proceedings of the 42nd Hawaii International Conference on System Sciences, p.1-11, January 05-08, 2009 [doi>10.1109/HICSS.2009.223]
[359] Preece, J., Rogers, Y., Sharp, H. (2002). “Interaction Design: Beyond Human-Computer Interaction”, New York: Wiley
[360] Basili, V.; Caldiera, G.; and Rombach, D. (1994). In: “The goal question metric approach” Export Find Similar, edited by: Marciniak, J., published by Encyclopedia of Software Engineering (1994).
[361] SAS (2010). “The Knowledge Exchange”. Retrieved from the web at August 2010. http://www.bettermanagement.com/topic/subject.aspx?f=11&s=704.
[362] Aaronson, S. (2010). “Statement on conceptual contributions in theory”. In: the blog of Scott Aaronson. Retrieved from the web a December 2011: http://www.scottaaronson.com/blog/?p=315